(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON
Managing Editor

HELEN M. DAVIS
Associate Editor

Editorial Board

GORDON ALLPORT
Harvard University

JOHN A. ANTOINETTI
Life Insurance Agency Management Association

ARTHUR G. BILLS
University of Cincinnati

E. A. BOTT
University of Toronto

LEONARD CARMICHAEL
Smithsonian Institution

HERBERT S. CONRAD
U. S. Office of Education

JOHN F. DASHIELL
Florida State University

LEONARD W. DOOB
Yale University

ROY M. DORCUS
University of California at Los Angeles

FRANK A. GELDARD
Princeton University

CLARENCE H. GRAHAM
Columbia University

ERNEST R. HILGARD
Stanford University

JOHN E. HORROCKS
Ohio State University

HEINRICH KLÜVER
University of Chicago

NORMAN L. MUNN
University of Adelaide, Australia

JOHN PAUL NAFE
Florida State University

WARREN H. TEICHNER
Northeastern University

LEE EDWARD TRAVIS
University of Southern California

ROBERT C. TRYON
University of California

JOSEPH ZUBIN
Psychiatric Institute, Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 64

1966

Copyright, 1966, by The Journal Press
Provincetown, Massachusetts, U. S. A.
All rights reserved

Published Bimonthly at 35 New St., Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts

AUTHOR INDEX

Albee, George W. 227
Allison, Roger B. 33
Asher, James J. 23
Baer, Daniel J. 63, 101, 249
Ball, Richard J. 241
Bartley, S. Howard 241
Beaulieu, C. 223
Brown, Donald W. 199
Cautela, Joseph R. 277
Chyatte, Conrad 233
Coxe, Malcolm 119
Davis, Joan 119
Dreger, Ralph Mason 161
Farley, Frank H. 113
Fisichelli, Vincent R. 119
Fox, David J. 265
Gengerelli, J. A. 159
Giannell, A. Steven 213
Gladstone, Roy 23
Haber, Audrey 119
Hobson, Sally 79
Jaffee, Cabot L. 127
James, W. T. 193
Jordon, Nehemiah 3
Karelitz, Samuel 119
King, H. E. 59
Korner, Ija N. 33
Lane, Ellen A. 227
McFarland, John 193
Norman, Ralph D. 49
Pervin, Lawrence A. 41
Pumroy, Donald K. 73
Ray, Wilbert S. 107
Rosenfeld, Laura 119
Rowntree, J. Timothy 151
Salzinger, Kurt 79
Salzinger, Suzanne 79
Schuh, Allen J. 69
Siegelman, Marvin 143
Solyom, L. 223
Steinmann, Anne 265
Stewart, Daniel K. 95
Taylor, Raymond G., Jr. 137
Tresselt, M. E. 91
Vestre, Norris 257
Vicory, Arthur C. 29
Wagner, Mahlon W. 151
Zwanziger, Max D. 33

TABLE OF CONTENTS

Experimenting with a *P-O-Q* unit: Complications in cognitive balance . . . 3
By Nehemiah Jordan

The significance of the extinction situation 23
By Roy Gladstone

The paired associate task as a predictor of foreign language fluency . . . 29
By Arthur C. Vicory and James J. Asher

Matching patients and therapists: A methodological approach 33
By Roger B. Allison, Jr., Ija N. Korner, and Max D. Zwanziger

Reality and nonreality in student expectations of college 41
By Lawrence A. Pervin

The interpersonal values of parents of achieving and nonachieving gifted children 49
By Ralph D. Norman

The retention of sensory experience: VI. Stimulus repetition and interference effects 59
By H. E. King

Scholastic aptitude and smoking attitude and behavior of college males . . 63
By Daniel J. Baer

A synthesis of personality theories by cluster analysis 69
By Allen J. Schuh

Maryland Parent Attitude Survey: A research instrument with social desirability controlled 73
By Donald K. Pumroy

Memory for verbal sequences as a function of their syntactical structure and the age of the recalling child 79
By Suzanne Salzinger, Kurt Salzinger, and Sally Hobson

A preliminary study on factors in learning in a how-to-study course . . . 91
By M. E. Tresselt

Communication and meaning 95
By Daniel K. Stewart

Height, weight, and ponderal index of college male smokers and nonsmokers 101
By Daniel J. Baer

Originality in problem solving as affected by single- *versus* multiple-solution train problems 107
By Wilbert S. Ray

Social desirability, extraversion, and neuroticism: A learning analysis . . 113
By Frank H. Farley

The phonetic content of the cries of normal infants and those with brain damage . 119
By Vincent R. Fisichelli, Malcolm Coxe, Laura Rosenfeld, Audrey Haber, Joan Davis, and Samuel Karelitz

The effect of short-term sensory deprivation on rote learning 127
By Cabot L. Jaffee

Racial stereotypes in young children 137
By Raymond G. Taylor, Jr.

Psychometric properties of the Wiggins and Winder Peer Nomination Inventory 143
By Marvin Siegelman

Methodology of relative sugar preferences in laboratory rats and deer mice 151
By Mahlon W. Wagner and J. Timothy Rowntree

Factors and factor analyses 159
By J. A. Gengerelli

Just how far can social change change personality? 161
By Ralph Mason Dreger

A study of form discrimination in the opossum 193
By W. T. James and John McFarland

The effects of anxiety, stress, and task difficulty on performance on a verbal learning task 199
By Donald W. Brown

Giannell's criminosynthesis theory applied to female homosexuality . . 213
By A. Steven Giannell

The effect of DNA on operant conditioning 223
By C. Beaulieu and L. Solyom

Comparative birth weights of schizophrenics and their siblings . . . 227
By Ellen A. Lane and George W. Albee

Brain blood-shift theory: Intersensory rivalry, time error, and closure illusions 233
By Conrad Chyatte

The study of target size in production of brightness, hue, and saturation for intermittent photic stimulation in the red end of the visible spectrum . 241
By S. Howard Bartley and Richard J. Ball

Sex differences in smoking attitude, behavior, and beliefs of college students 249
By Daniel J. Baer

The effects of phenothiazine drugs on verbal conditioning of schizophrenics 257
By Norris Vestre

Male-female perceptions of the female role in the United States . . . 265
By Anne Steinmann and David J. Fox

Desensitization factors in the hypnotic treatment of phobias 277
By Joseph R. Cautela

$30.00 per annum
$15.00 per volume
$7.50 per half volume

BIMONTHLY
Three volumes per year
Immediate publication

September, 1966
Volume 64, First Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

SEPTEMBER, 1966

(Manuscripts are printed in the order of final acceptance)

Experimenting with a *P-O-Q* unit: Complications in cognitive balance . . . 3
By Nehemiah Jordan

The significance of the extinction situation 23
By Roy Gladstone

The paired associate task as a predictor of foreign-language fluency . . . 29
By Arthur C. Vicory and James J. Asher

Matching patients and therapists: A methodological approach 33
By Roger B. Allison, Jr., Ija N. Korner, and Max D. Zwanziger

Reality and nonreality in student expectations of college 41
By Lawrence A. Pervin

(OVER)

Copyright © 1966 by The Journal Press
Provincetown, Massachusetts, U. S. A.
All rights reserved
Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts

The interpersonal values of parents of achieving and nonachieving gifted children 49
By Ralph D. Norman

The retention of sensory experience: VI. Stimulus repetition and interference effects 59
By H. E. King

Scholastic aptitude and smoking attitude and behavior of college males . . 63
By Daniel J. Baer

A synthesis of personality theories by cluster analysis 69
By Allen J. Schuh

Maryland Parent Attitude Survey: A research instrument with social desirability controlled 73
By Donald K. Pumroy

Memory for verbal sequences as a function of their syntactical structure and the age of the recalling child 79
By Suzanne Salzinger, Kurt Salzinger, and Sally Hobson

A preliminary study on factors in learning in a how-to-study course . . 91
By M. E. Tresselt

Communication and meaning 95
By Daniel K. Stewart

Height, weight, and ponderal index of college male smokers and nonsmokers . 101
By Daniel J. Baer

Originality in problem solving as affected by single- *versus* multiple-solution training problems 107
By Wilbert S. Ray

Social desirability, extraversion, and neuroticism: A learning analysis . . 113
By Frank H. Farley

The phonetic content of the cries of normal infants and those with brain damage 119
By Vincent R. Fisichelli, Malcolm Coxe, Laura Rosenfeld, Audrey Haber, Joan Davis, and Samuel Karelitz

The effect of short-term sensory deprivation on rote learning 127
By Cabot L. Jaffee

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 3-22.

EXPERIMENTING WITH A *P-O-Q* UNIT: COMPLICATIONS IN COGNITIVE BALANCE*[1]

Institute for Defense Analyses, Arlington, Virginia

NEHEMIAH JORDAN

A. INTRODUCTION

The *p-o-x* unit has been playing a minor but noticeable role in the professional social-psychological literature (psychology as distinguished from sociology) for several years. The unit is derived from a short paper published by Heider (2) in 1946. Heider points out that there seems to be a harmony between the attitudes that a person has towards entities (be they human or nonhuman) and events in his world, and the objective relations that the person perceives to hold both between those entities as such and with respect to himself. This harmony can be expressed by the almost truistic statement that "we like people we like to like the things we like," and that one experiences a degree of discomfort when this is not the case. Another truistic aspect of this harmony is the almost universal expectation that good men do good acts and bad men do bad ones; when we are confronted with what we think to be a good act done by a bad man, or a bad act done by a good man, the harmony is disturbed and the situation becomes problematical. We then seek explanations. The most general explanation that restores the harmony imputes an ulterior "bad" motive to the acts of the bad man and a vision of some superordinate good underlying the act of the good man. It can easily be seen that this is the stuff of which rationalizations are made.

This harmony, at present, is called "cognitive balance."

Heider speculated: To what extent does balance and imbalance, as such, balance and imbalance *independent* of the actual substance constituting the perceived situations, affect behavior? Can it be said that, in addition to or in conjunction with whatever else determines human behavior, the mere distribution of attitudes in a given cognitive organization or structure has a noticeable and important effect upon behavior? Heider thought this to be the case. If so, then the effect upon behavior must also be independent of a practical

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 18, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Previously published as Internal Note N-296(R), Institute for Defense Analyses, Economic and Political Studies Division, Arlington, Virginia, December, 1965.

goal achievement or of reward and punishment considerations. The presumed effects of balance move consequently into the realm of esthetics, ethics, and propriety—a realm where actions are essentially undertaken for their own sake and not for any ulterior purpose.

In his discussion of balance, Heider successfully represented the complexities of cognitive balance by means of a delusively simple schema, a schema that seemed to lend itself to relatively easy experimental testing. Thereby a challenge emerged. It is well known that the realm of esthetics, ethics, and propriety is very difficult to submit to experimental investigation. A few hardy souls have tried such investigations, but upon reading their published reports an inescapable feeling emerges that somehow or other the essence of the matter has been lost in the shuffle. If cognitive balance is truly within this important but experimentally elusive realm of values, and if one can experiment with it successfully, may this not be a way for achieving some more general experimental understanding of the entire realm proper? A research problem crystallized: can it be demonstrated experimentally that balance and imbalance, *as such*, can affect behavior? Or, to enunciate Heider's speculations more accurately, can it be demonstrated experimentally that balance, *as such*, is a more preferred, a more "proper" state than imbalance?

Before going on to the experiments, some words are in order to sketch out Heider's schema in discussing these problems. The research problem is the effect of the perception or cognition of a balanced or imbalanced configuration, *as such*, upon a person. Let the person being studied be denoted by the symbol p; the person can also be called "the actor" or "the subject." The other two elements playing an important role in this problem area are (what is currently nebulously called) "attitudes" and "entities," be they human or nonhuman, animate or inanimate. Attitudes relate to evaluations, be they positive or negative. Let all these evaluations—which can range from a mild liking or disliking, a mere favoring or disfavoring to an extreme acceptance or rejection on an ethical or religious nature—be denoted by the symbols $+L$ and $-L$ respectively. Let the animate entities, other people, be denoted by symbols such as o, q, r, etc., and the inanimate entities by x, y, z, etc. Abstract configurations constructed by means of these symbols (with one more soon to be introduced) should therefore be experienced as being proper or improper, desirable or undesirable, pleasant or unpleasant, etc., *as such,* necessarily independent of specific context.

Here are some examples of how the symbols can be used. Obviously, it is balanced that a person esteems himself: $p + Lp$. If a person likes himself he should also have nice attributes. If any positive attribute is denoted by x, then it is balanced for the person who likes himself to have that attribute: $p + Lp$; $p + Lx$; p is x. Similarly, it is balanced when a person perceives that anything (x) caused by him is good: $p + Lp$; p caused x; $p + Lx$. The person's perception that something caused by him is bad seems intuitively to carry with it an element of unpleasantness or of discomfort: $p + Lp$; p caused x; $p - Lx$. A subtler form of imbalance: A person perceiving a good action that he could have taken but did not: $p + Lp$; p did not cause x; $p + Lx$. It is of course balanced if the person considers the action to be bad: $p + Lp$; p did not cause x; $p - Lx$. A final example: Suppose a third person comes to the subject and tells him that he is similar to another person, o. This is balanced if the subject likes, values, esteems, etc., the other person: $p + Lp$; $p + Lo$; p is similar to o. It is imbalanced if the subject dislikes the other person: $p + Lp$; $p - Lo$; p is similar to o.

There are several "nonsymbolized" relations in the preceding paragraph —*is, caused, similar*—that have nothing to do with attitudes; yet they affect balance. They are possible structural relations between the entities constituting the perceived situations. Let all such possible structural relations be denoted by the symbol U.[2] With this symbol all the above balanced and imbalanced situations can be subsumed under two symbolic expressions: $p + Lp$; $p + Lx$ (or o); $p + Ux$ (or o), and $p + Lp$; $p + Lx$ (or o); $p - Ux$ (or o).

The situation does not change when we implicitly assume the person's self-esteem (after all, who dislikes oneself?) and consider situations with another person. It is intuitively evident that the situation *friends sharing the same attitudes* is balanced: $p + Lo$; $p + Lx$; $o + Lx$ and $p + Lo$; $p - Lx$; $o - Lx$. It seems just as intuitively evident that *"enemies" not sharing the same attitude* is also balanced: $p - Lo$; $p + Lx$; $o - Lx$ (something one likes becomes "contaminated" to a certain degree with the perception that one's enemy likes it too); similarly, an enemy liking a disliked thing is balanced: $p - Lo$; $p - Lx$; $o + Lx$ (one should recall what happened to Wagner's music in the United States when it became known that Hitler liked it). This also holds for U relations. It seems to be pleasant and proper for our friends to have valued attributes: $p + Lo$; $p + Lx$; $o + Ux$; and it is just as proper for enemies to have negative attributes (villains are rarely portrayed as truly handsome, truly smart, etc.): $p - Lo$; $p - Lx$; $o + Ux$. Actions taken by one's friends should be good, while actions taken by one's enemies "naturally" should be bad. Friends should be similar to "good guys"; enemies, to "bad guys," etc. Other people, *q, r,* etc., and entities *y, z,* etc., can be introduced, but then the situations become complex and difficult to follow through.[3]

---

[2] This is taken from the term "unit-formation." The assumption underlying the choice of this symbol comes from Gestalt psychology. Just as the visual field of a person consists of organized wholes, units, whose organization can be explained by the Gestalt laws of unit formation (8, chap. IV), so the perception or cognition of a social situation is organized. U relations, as well as others, obviously, play their role in the organization of cognized social units.

B. Early Experiments

For contingent, theoretically irrelevant reasons, the first systematic research to be published (5) attempting to test the causal efficacy on behavior of balance and imbalance, as such, restricted itself to imagined social situations containing the subject (p), another person (o), and an impersonal entity (x). With this the *p-o-x* unit was inadvertently born. There is nothing intrinsically holy or significant about the *p-o-x* unit; it seemed at the time and subsequently has proven itself to be a simple and easy way of investigating cognitive balance.

Given a *p-o-x* schema that consists of three related pairs (p and o, p and x, and o and x) and four possible relations ($+L$, $-L$, $+U$, and $-U$), 64 discrete units are constructable. In his original speculations Heider excluded 15 of these from consideration; they were the units containing three U relations, independent of sign, and units containing three negative relations,[4] be it L or U. This is a fine and, as later experimentation has demonstrated (6), valid distinction.

Subjects were presented with abstract descriptions of these 64 possible *p-o-x* units, asked to imagine themselves as being p in them, and then asked to rate the imagined situations as to experienced pleasantness and unpleasantness. The description of the situations was as abstract as possible, an almost word-for-symbol transformation consistent with meaningful communication to the subjects. It was hoped that this abstract presentation would conduce the subjects to attend to balance and imbalance *as such* and not attend to matters extraneous to the experimenter's interests. This was not successful; balance and imbalance did have an effect upon the subject's ratings but in an unexpected way that is difficult to interpret as "their being attended to."

The experimental results in general did not live up to the experimenter's expectations. There was only one way to interpret legitimately the data obtained: contrary to the hypotheses, exactly 50 per cent of the balanced situations were experienced as being basically unpleasant, whereas 50 per cent of the imbalanced situations were experienced as being slightly pleasant. As already mentioned, balance and imbalance showed an unexpected effect: the balanced situations, both pleasant and unpleasant, showed a far greater sensitivity to the gamut of experimental manipulations—of which there were quite a few—than did the imbalanced situations, whether pleasant or unpleasant. This was interpreted as demonstrating that balance is a strong Gestalt, while imbalance is a weak Gestalt. With this, Heider's "hypothesis emerged unscathed." After all, what was argued was that balance and imbalance have an effect, and a striking effect was found. Actually, this new effect was far more in line with the desires of the experimenter than was the so-called predicted effect. By being able to develop an argument linking balance with a strong Gestalt and involving Wertheimer's law of Pregnanz (8, p. 110 *ff.*): a strong Gestalt is more "good" than a weak Gestalt, a psychological force towards balance is, *ipso facto,* demonstrated; moreover, it is a force rather intimately related to the areas of values, oughts, and ethics.

---

[3] As can be seen, Heider's scheme is a rather loose borrowing from the symbolic logic. Cartwright and Harary (1) have rewritten the above considerations in a rigorous, mathematical, point-graph theory notation. Their resultant mathematical formulae are elegant; it admittedly is a *tour de force;* but it has yet to be demonstrated that the mathematical elegance and ingenuity contribute much of substance to our knowledge in this area.

[4] In his later thinking, Heider considers situations with three negative relations as constituting some weak, less preferred form of balance (3, page 206).

The data had another fascinating aspect that was noted at the time but the significance of which was not appreciated. The data were submitted to a rather complex analysis of variance, each one of the possible 64 *p-o-x* units serving as an independent "treatment." The total effect of the independent variable was expressed by the sum of squares it contributed to the total sum of squares of the experiment. The independent variable sum of squares could, in turn, be partitioned into 63 orthogonal partitions (10 main effects and 53 interaction effects). Subsequent analysis showed that 19 orthogonal partitions, *all psychologically meaningful*, accounted for almost 92 per cent of the total treatment sum of squares. This has rarely, if ever, been the case in experimentation of attitudes and cognition. It could reasonably be argued that, "for all practical purposes," the effects of the independent variable, the 64 imagined situations, on the ratings could be completely explained in meaningful psychological terms.

Soon afterwards another experiment was conducted along similar lines (4). In this experiment what amounted to 18 situations of the preceding experiment were further expanded by the addition of two new types of variation not found earlier: two degrees of attitude strength (like very much and like moderately, dislike very much and dislike moderately) between *p* and *x*, and *o* and *x*, the attitude between *p* and *o* being kept constant (like very much, dislike very much); and two types of inanimate entities (*x*):

things and values. In all other respects this experiment was very similar to the preceding experiment. It was possible to rank-order correlate the data of this experiment with the average ratings of the 18 situations in the preceding experiment; the rank-order correlation is .985. That high a correspondence between data of two independent experiments is also a rarity in research on attitudes and cognition.

Sixty-four of the 120 situations presented to subjects in this experiment could be subjected to an analysis of variance similar to the one in the preceding experiment. Twenty orthogonal partitions were identified that accounted for 96.5 per cent of the total sum of squares generated by the 64 situations.

Excitement emerged; here was something significantly novel in psychological research: an underlying invariance embedded in a richly varying pattern of individual ratings. It can be argued that one of the main reasons dooming contemporary psychological research to functional sterility is that it has never found a base "solid" enough for it to develop a "feel" for the causal nexus underlying psychological phenomena. Almost all complex measurements of psychological phenomena rarely go beyond mere statistical significance; all measurements leave vast amounts of variance unaccounted for. This unaccounted variance "floats freely" in the world of the research psychologist and has an effect upon him similar to the effect of free-floating anxiety upon an average human being. We stand with both our legs firmly implanted upon solid air; nothing substantive can be built; no continuity of research can be maintained; no understanding can be reached. In order to justify ourselves to ourselves and to hide our impotence and nakedness from the prying eyes of the world at large, we surround ourselves with a web of words, place all our faith on that conceptual morass called "scientific method," and long ardently for that never-never land of mathematical models, that nirvana where all problems will cease.

What can be done? What is needed? What seems to be needed is an area, any area, of psychological phenomena that can be studied in depth, that can be submitted to controlled manipulation and change, yet that will demonstrate a patterned invariance or order through change; only through such a study can a feel for psychological causality be developed; only through such a study can laws of behavior, if there be such, be ultimately reached. After all, the road leading to the discovery of the laws of motion of the solar system started with careful observation and analyses of swinging pendula and rolling stones. We pride ourselves on using the "advanced" physical sciences as a model; why do we not choose the correct model?

Could it be that in this research area of changes in people's markings on a scale that are a function of changes in what they are told to imagine—changes that are very sensitive to changes in instructions, yet at the same time exhibit a surprising invariance, a surprising stability (for what else can the high correlation between experiments and the high accounting for the sum of squares within the experiments indicate?)—could it be that in this area of psychological research the road will be found to the essential psychological order that is the necessary and sufficient condition for creating a science of human behavior?

C. COMPARISONS WITH PREVIOUS EXPERIMENTS

Data from old forgotten experiments were dug out and compared with the data from the two experiments just discussed. A pattern did emerge. Both the pattern and its detailed justification are presented in a monograph (7); a cursory summary of this pattern is sufficient for present purposes. Two factors seemed to play the major role in determining the subjects' ratings of the imagined *p-o-x* situations: (*a*) the relation between the two actors, *p* and *o*; and (*b*) "what the future seemed to have in store" for the situation. Balanced situations with a positive relation between the actors—i.e., balanced situations with friends—were all judged to be highly pleasant. The formal conditions defined as being balanced "reinforce" the positive relation between *p* and *o*, since they involved common attitudes towards the impersonal entity. Hence, the situation was stable, and the friendship was not threatened. Balanced situations with a negative relation between the actors—i.e., balanced situations with "enemies"—were all judged to be highly unpleasant. The formal conditions defined as balance "reinforce" this "enmity," since they dictate a conflict of attitudes with respect to the impersonal entity. Imbalanced situations with friends were rated as being highly unpleasant as well: there is a conflict of attitudes among friends; the future fate of the friendship is in doubt, the friendship is threatened. Imbalance creates an unstable situation. And, finally, the imbalanced situations with enemies were rated as being *mildly unpleasant*. Formal conditions defining balance dictate a common attitude towards the impersonal entity in such circumstances. But *o* cannot be much of an imagined enemy if he shares a common attitude with *p* towards *x*, can he? Even more, may not this common attitude become a vehicle whereby the enmity will, at least, be dissipated if not turned into friendship. Again imbalance creates instability but, in this case, a desired instability that bides a welcome change.

With this analysis, balance and imbalance seemed all but to disappear; they play no explicit role in the preceding "explanation" of the subjects' behavior. There are three discrete ways to treat this, each being true to some extent or other. It can be argued that despite the abstract presentation of the *p-o-x* situations, the experiment was unsuccessful in getting the subjects to focus upon balance *as such*. Rather the subjects "perceived" the relations as such and reacted to them instead. Hence, the experiment was invalid as far as testing the effects of balance, *as such*. The problems connected with the validity of psychological tests are well known; it is surprising, however, how little attention has been placed on the validity of psychological experiments. All too often this problem has been avoided by parading the dazzling slogan of "operational definition." But operational definitions *per se* do not solve the problem. Everybody would object were someone operationally to define intelligence as the number of times a person scratches his nose during a finite measurable period of time; may it not be the case that the operational definition of the effects of balance assumed in these experiments, the kind of marks made by subjects on a special piece of paper, is of the same order? Experiments and operational definitions may be just as invalid for their specific purposes as are many tests.

On the other hand it can be argued that the *p-o-x* unit is an instantiation of the abstract concept: balance/imbalance. As is the case of every instantiation of an abstract concept, it itself can be discussed and analyzed in terms of its concrete elements without ever having to introduce the abstract concept underlying it. In the present experiments it is the pattern of relations that "defines" balance and imbalance and when the subjects' marks on the paper are explained in terms of these relations, they are *ipso facto* also explained in terms of balance and imbalance as well.

Finally, an effect univocally related to balance and imbalance and nothing else has been identified. As mentioned above, the data of the first experiment showed that balanced situations were far more sensitive to experimental changes of the independent variable than were the imbalanced situations. The same was found in the data yielded by the second experiment. It was possible to compare the effects of 82 comparable experimental changes under conditions of balance and imbalance (6). The standard deviation of the array of these 82 effects under conditions of balance is 15.61; under conditions of imbalance it is 1.48. An effect is an effect, even though it was not a predicted effect.

And so matters stood for quite a few years. The analysis of four experiments on the *p-o-x* unit yielded great excitement and great expectations. But little valuable research was undertaken.

D. NEW RESEARCH

Recently a new paper by Price, Harburg, and McLeod (9) was published, the data of which are relevant and comparable to the data already discussed. They are disturbing; they have dampened the excitement; they do not fit. It is one of the main purposes of this paper to present these data in detail to explicate their implications, to indicate clearly the nonfittedness. Problems will be raised but not answered. The nonanswering of problems is not due to a lack of will, but to a lack of ability. "Heroic" analytic attempts were undertaken, to try to integrate these data to the data of the preceding experiments, to no avail, but disconfirming data are just as important as those that confirm, and should be given an equal "spread."[5]

In many respects the new experiment is similar to the two preceding ones—hence, the comparability of the data. Again subjects were asked to imagine themselves in described situations and to rate the imagined affect of those situations on a scale similar to the one used in the previous experiments. But there are several differences, at least some of which are probably significant and which may, ultimately, explain the discrepancies that are at present disturbing.

What will probably turn out to be the most significant difference is that the situations presented to the subjects were not *p-o-x* units but *p-o-q* units: i.e., the third element of the situations was not an inanimate entity (x) but another person (q). Now this may introduce all kinds of (hopefully, merely temporary) unanticipated consequences. It will be recalled that the *p-o-x* unit was originally chosen because of an intuitive feeling that it was simple and would enable the subjects to attend to balance and imbalance as such. Later research disclosed that the *p-o-x* unit was not that simple, and analysis showed that subjects were affected by all kinds of subtle complexities. Yet, when all is said and done, a dyad is basically a stable social grouping. On the other hand, a triad is a notoriously unstable social grouping giving rise to all kinds of currents and crosscurrents, intrigues and conflicts; who can tell the additional complexities added to the imagined situation when x is replaced by q?

Another difference, which may also turn out to be significant, is that these *p-o-q* situations were not abstract, general, but were concrete, specific. The subjects were asked to name people they knew at the time who fit the relationships prescribed in the experimental *p-o-x* situations. Each subject was then asked to rate imagined situations containing those names. A person whom one actually knows is not an abstract o or q. He is a complex bundle of relations, connections, and psychological implications with all kinds of other entities and aspects of the commonly shared environment. Again it is difficult to know what complications are introduced thereby.

---

[5] For a critique of the Price *et al.* experiment see Jordan (6).

A third difference, seemingly small but worth mentioning, is the difference in scale used for measuring the attitude strength. Formally it was the same as the scales used in the preceding two experiments: a graphic rating scale with three given points—the midpoint and the two end points. But the end points were denoted differently; in the preceding experiments the end points were denoted most pleasant and most unpleasant, respectively; in the present experiment they were denoted uneasy and pleasant. Whereas the earlier scale truly did represent a closed, complete, unidimensional continuum, ranging through all possible values for that continuum, the later scale is neither complete, closed, nor truly a continuum—"uneasy" is not an antonym to "pleasant." In the business of attitude measurement, slight, seemingly irrelevant changes in the measuring instrument may trigger off significant changes in subjects' behavior. It is not known whether such was the case in the present instance.

The relatively novel independent variable in this experiment was the variation of the strength of attitudes held between the three elements constituting the situation. Hershkowitz (4) used two degrees of attitude strength between p and x and o and x; in the present experiment three degrees of attitude strength between the three pairs were used in many cases. Unfortunately, the choice of situations to be presented to the subjects was not systematic; no rationale for the choice is given by the experimenters nor can one be educed from the choice. Had the selection been systematic and relatively complete, the data would have been far more "powerful."

The basic "nature" of such situations (i.e., whether they are balanced or imbalanced, etc.) is determined by the signs of the attitudes between the three constituting dyads and not by the intensity of the attitudes; hence, it is proper to speak of basic p-o-x units when the discussion is restricted to the signs. The 58 specific situations presented to the subjects in the experiment under discussion were all generated from four basic situations: 1) $p + Lo$; $p + Lq$; $o + Lq$. 2) $p + Lo$; $p + Lq$; $o - Lq$. 3) $p + Lo$; $p - Lq$; $o + Lq$. 4) $p + Lo$; $p - Lq$; $o - Lq$. Since three independent strengths of attitude could be assigned to each attitude in each basic situation, 27 discrete experimental situations can be generated from each basic situation. This was not done, however. Twenty-five variations were chosen from the first basic situation, 24 from the second, six from the third, and three from the fourth. Hence, the data are incomplete from a systematic standpoint, and there are displeasing loose ends; it would be very nice were the experiment to be repeated with a systematic choice of experimental situations.

The data are presented in Table 1. Table 1 consists of five columns. The first column gives the 27 possible variations of attitude strengths for each basic situation. The next four columns give the mean ratings and standard deviations of those means for the various experimental situations; cells without entries indicate possible situations that were not presented to subjects for rating. The means are an average of 100 individual ratings. The mean scores in Table 1 are a linear transformation of the mean scores given by Price *et al.* (9). The lower the score, the more preferred the situation is rated to be. To obtain the original scores, subtract 55 from each transformed score and multiply the remainder by —1.

As a first step in the analysis of the data, it is interesting to see how they

TABLE 1
THE MEAN RATINGS AND STANDARD DEVIATIONS OF THE *P-O-Q* SITUATIONS

	+*L*, +*L*, +*L*		+*L*, +*L*, —*L*		+*L*, —*L*, +*L*		+*L*, —*L*, —*L*	
Attitude strength	Mean rating	*SD*	Mean rating	*SD*	Mean rating	*SD*	Mean rating	*SD*
1, 3, 3	40	19	69	20	60	18	43	18
2, 3, 3	32	18	76	17	—	—	—	—
3, 3, 3	28	18	85	14	72	19	34	17
1, 3, 2	47	15	64	18	59	15	—	—
2, 3, 2	—	—	70	13	—	—	—	—
3, 3, 2	54	17	71	15	69	15	—	—
1, 3, 1	48	13	65	12	58	13	—	—
2, 3, 1	42	15	71	13	—	—	—	—
3, 3, 1	49	17	76	12	63	14	43	15
1, 2, 3	42	19	69	17	—	—	—	—
2, 2, 3	—	—	70	18	—	—	—	—
3, 2, 3	37	17	76	15	—	—	—	—
1, 2, 2	46	13	62	12	—	—	—	—
2, 2, 2	43	15	63	9	—	—	—	—
3, 2, 2	39	13	68	13	—	—	—	—
1, 2, 1	45	11	63	11	—	—	—	—
2, 2, 1	46	16	66	11	—	—	—	—
3, 2, 1	46	13	67	13	—	—	—	—
1, 1, 3	43	18	60	16	—	—	—	—
2, 1, 3	54	17	—	—	—	—	—	—
3, 1, 3	56	20	56	19	—	—	—	—
1, 1, 2	45	14	59	12	—	—	—	—
2, 1, 2	52	12	—	—	—	—	—	—
3, 1, 2	48	15	56	14	—	—	—	—
1, 1, 1	46	15	59	10	—	—	—	—
2, 1, 1	48	14	—	—	—	—	—	—
3, 1, 1	47	15	53	14	—	—	—	—

Note: The data in Table 1 are adapted from "Positive and negative affect as a function of perceived discrepancy in AB*X* situations" by Price, Harburg, and McLeod (9), and are used by permission of *Human Relations* and the senior author.

relate to the data of the earlier experiments; a comparison is possible. As mentioned, Hershkowitz also varied the strength of attitudes in his experiment. The variation, however, was not as extensive as in the present experiment; the attitude between *p* and *o* was constant—likes or dislikes very much: i.e., of a magnitude 3—and there were two degrees of attitude strength between the two actors, *p* and *o*, and the impersonal entity, *x*—like or dislike very much and like or dislike moderately: i.e., magnitudes 3 and 2. Eleven situations in the Price *et al.* experiment are directly comparable to 11 situations in the Hershkowitz experiment as far as the distribution of attitude strength is concerned. A rank order correlation can be calculated between these two sets of 11 situations. The relevant data are presented in Table 2. Note that two means are given for the Hershkowitz experiment: one mean for situations where *x* is a value and another for situations where *x* is an object.

TABLE 2
A COMPARISON OF DATA FROM THE PRICE *et al.* EXPERIMENT AND THE HERSHKOWITZ EXPERIMENT

	Price *et al.* experiment		Hershkowitz experiment			
			Value		Object	
Situation	Rating	Rank	Rating	Rank	Rating	Rank
$3 + L;\ 3 + L;\ 3 + L$	28	1	2.6	1	4.2	1
$3 + L;\ 3 + L;\ 3 - L$	85	11	81.8	11	80.4	11
$3 + L;\ 3 - L;\ 3 + L$	72	9	80.2	10	59.0	8
$3 + L;\ 3 - L;\ 3 - L$	34	2	12.8	3	11.0	2
$3 + L;\ 3 + L;\ 2 + L$	54	5	42.0	5	24.2	4
$3 + L;\ 3 + L;\ 2 - L$	71	8	66.0	9	57.8	7
$3 + L;\ 3 - L;\ 2 + L$	69	7	63.6	7	60.8	9
$3 + L;\ 2 + L;\ 3 + L$	37	3	23.9	4	36.0	5
$3 + L;\ 2 + L;\ 3 - L$	76	10	61.4	6	65.5	10
$3 + L;\ 2 + L;\ 2 + L$	39	4	11.4	2	15.8	3
$3 + L;\ 2 + L;\ 2 - L$	68	6	63.7	8	54.2	6

Note: The data in Table 2 are adapted from "Positive and negative affect as a function of perceived discrepancy in *ABX* situations" by Price, Harburg, and McLeod (9), and are used by permission of *Human Relations* and the senior author; and from "Interpersonal agreement and disagreement on objects and values—A preliminary study" by Hershkowitz (4), and are used by permission of the author.

Calculation yields a rank order correlation coefficient of .873 between the Price *et al.* experiment with the Hershkowitz experiment (value), and .945 with the Hershkowitz experiment (object). The first correlation coefficient is disappointingly low;[6] the second is in the mid-90s like the others. Nevertheless, they are both impressively high, as these things go. It is interesting to note that the correlation with the situations containing *x* (objects) is higher than the correlation with the situations containing *x* (values). Should the conclusion be jumped to that contemporary college students treat other people more like objects to be manipulated and used rather than entities of intrinsic value as such?

---

[6] With respect to the low correlation, inspection shows that the situation $3 + L;\ 2 + L;\ 3 - L$ (value) yields more than half of the sum of the deviations squared. Were this situation to be eliminated and the array reduced to 10 ranks, the correlation jumps to .945. This is much more pleasing. No convincing argument can be brought forth to justify such a procedure, but it is one that is often found in the literature and seems therefore justified by public sanction.

This relatively high correlation indicates that the psychological explanations that seemed to account for the subjects' behavior in the earlier experiment should also hold for this last experiment. This must be the case for the 11 situations that are almost direct counterparts with the "Hershkowitz" 11 situations. But there are an additional 47 situations that are not comparable to any of the situations in the Hershkowitz experiment—what about them?

Before exploring this last question, it may be valuable to review the psychological principles used to explain the subjects' behavior in the earlier experiments. It was asserted that the subjects seemed to take two things "into account" in "determining" their ratings of the situations: (*a*) the relation between *p* and *o*, and (*b*) what the future seemed to hold in store: i.e., whether the situation was stable and resistant to change (balanced) or not stable and relatively easily changed. Given a positive relation between *p* and *o*—i.e., friendship—and given stable situations where their attitudes towards *x* agree so that no conflict threatens to "mar" the future, the situations are rated as being quite pleasant. However, given an unstable situation with a "conflict" of attitudes towards *x*, the future is threatening and the situation is rated as quite unpleasant. The state of enmity, whether real or imagined, is unpleasant; people do not like the idea of enemies.[7] In a stable situation the enmity is "reinforced" by the conflict of attitudes and the future bears little hope for change; in the Hershkowitz experiment such situations are rated as being the most unpleasant. In an unstable situation, the "enemies" share a common attitude towards *x;* this does not bide too badly for the future. Nevertheless, enmity is enmity and these situations are rated as being unpleasant too, albeit not too unpleasant. Differences in attitude intensity between the actors towards *x* have an effect similar to the effect of formal imbalance; they make the situation decreasingly stable. The effect, however, is weaker than that of formal balance and was therefore called an effect of secondary balance or imbalance.

The terms "into account" and "determining" were put in quotation marks at the top of the preceding paragraph. If these terms are taken literally, the impression may be gained that the subjects consciously analyzed the situations, taking them into explicit account. This may have been the case, but the present writer doubts it. More probably the subjects were just aware of an "unanalyzed" feeling that they then proceeded to mark off on the scale. It was only subsequent analysis of the changes in marks with the changes in situations that led the analyst to conclude that these factors played the role they did, it being irrelevant whether the subjects were aware of them or not.

---

[7] Note that this is another form of Heiderian balance.

Immediate inspection of the data with respect to situations of the form $+L; +L; +L$ and $+L; +L; -L$[8] entered in Table 1 above shows that no simple order, if there is one, is apparent. To look for a possible order, one must reorganize the data. The reorganization is found in Figure 1.

The various situations are organized as to their ratings and "forms." What was sought was an inner order determined by the difference in ratings with

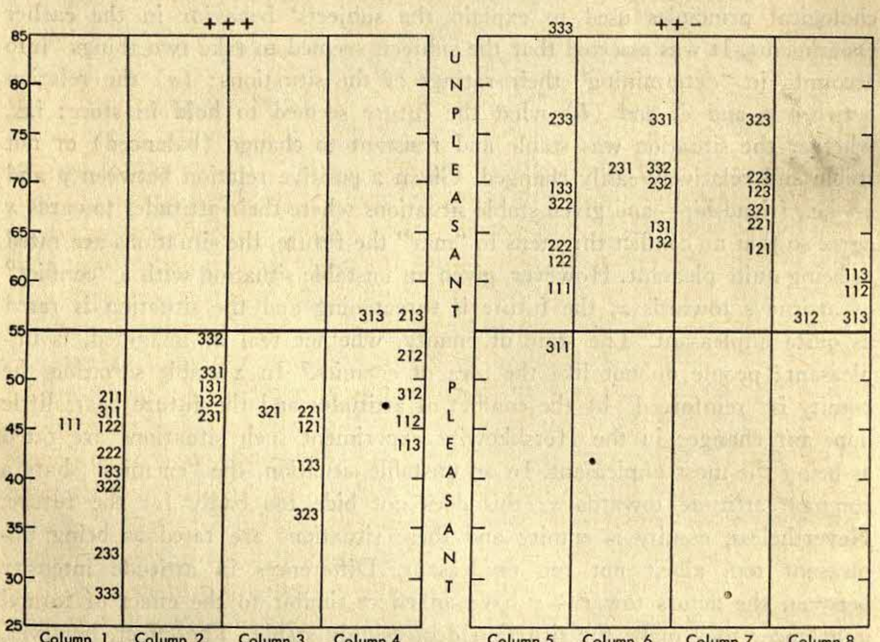

FIGURE 1

DISTRIBUTION OF *P-O-Q* SITUATIONS ACCORDING TO FORM AND MEAN RATING

The data are adapted from "Positive and negative affect as a function of perceived discrepancy in AB*X* situations" by Price, Harburg, and McLeod (9), and are used by permission of *Human Relations* and the senior author.

---

[8] **There are too few experimental variations of the other two forms for a rewarding analysis, and they will not be considered in the subsequent discussion.**

respect to individual forms of situations. It was felt that these forms are Gestalten of sorts, the implication being that identical changes of "independent variables" (i.e., the elements constituting the form) need not have the same effect in different forms. Although the "last word" on this has yet to be said, the order found in Figure 1 vindicates the feeling.[9] First the data were segregated as to their basic situational forms, the balanced situation $+L$; $+L$; $+L$ and the imbalanced situation $+L$; $+L$; $-L$. The data were then further aggregated into four groups. In the first group, columns 1 and 5, the attitude intensities of the actors towards *q* is the same. In the other three groups they differ; they are various instances of secondary imbalance. In columns 2 and 6, *p* has a strong attitude towards *q*, while *o* has various degrees of weaker attitudes. In columns 3 and 7, *p* has a mild attitude towards *q*, while *o* has either a strong or a slight attitude. And in columns 4 and 8, *p* has a slight attitude towards *q*, while *o* has various degrees of stronger attitudes.

The power of the psychological explanation in the earlier experiments was that once pointed out it was "self-evident," in most cases. Similar self-evidence can be found in the data entered in columns 1 and 5. Let column 5 be considered first. The situations in this column consist of two basic aspects: various degrees of strength of friendship and various degrees of conflict of *equal* attitude intensities between the actors with respect to *q*. It is self-evident that threats to friendship are unpleasant and conflicts, as such, are unpleasant. It is just as self-evident that the stronger the friendship threatened and the stronger the conflict, the more the experienced unpleasantness. Hence, as column 5 shows, a situation with friends with a strong bond [3] and a strong conflict [3, 3] is rated as highly unpleasant. For the next five situations in this column, the unpleasantness decreases first as the bond between the friends weakens and then as the intensity of the conflict weakens: i.e., the latter plays a stronger, more potent role in determining the unpleasantness rating than the former. No self-evident nor conceptual justification for this difference in potency is available; it is an empirical finding in the present instance. The order breaks down however when the attitude intensity of the conflict becomes slight [1, 1]; the situation $+3$; $+1$; -1 actually crosses over into the "pleasant region." It is argued in the present context that so mild a conflict has a different effect from that of the stronger conflicts above, though it still has an effect. The data in column 1, though not as neat, present a similar pattern: 3, 3, 3 is preferred to 2, 3, 3 and both are preferred to 1, 3, 3; 3, 2, 2 is preferred to 2, 2, 2 and both are preferred to 1, 2, 2; and again one finds a breakdown of order in the form —1, 1, 1.

---

[9] It is interesting to note that the forms found in the table came to mind intuitively as the proper mode for organizing the data. Attempts were later made to organize the data systematically in other ways, according to other forms. No other way yielded as neat an ordering of the data. There is room for intuition in science.

But with this "self-evident" order ends. Despite the ability to identify orders in the remaining columns, they are not self-evidently related to the psychological explanation—if they are related at all. If one ranks the situations in columns 6 and 7 ordinally, ranking ties as one wishes, the arrays given in Table 3 are obtained.

TABLE 3
RANKED SITUATIONS IN COLUMNS 6 AND 7 OF FIGURE 1

Column 6	Column 7
3, 3, 1	3, 2, 3
3, 3, 2	2, 2, 3
2, 3, 1	1, 2, 3
2, 3, 2	3, 2, 1
1, 3, 1	2, 2, 1
1, 3, 2	1, 2, 1

Note: The data in Table 3 are adapted from "Positive and negative affect as a function of perceived discrepancy in AB*X* situations" by Price, Harburg, and McLeod (9), and are used by permission of *Human Relations* and the senior author.

Each column yields a fascinating order but they (*a*) contradict each other and (*b*) do not make much sense in terms of the principles. In column 6 the attitude between *p* and *o* plays the major role (is most potent) in determining the unpleasantness of the situation, whereas in column 7 the conflict in attitude intensity with respect to *q* plays the major role; why should 3, 2, 1 be preferred to 1, 2, 3 in column 7? Intuitively a weaker conflict between strong friends should be more unpleasant than a stronger conflict between slight acquaintances—for what else could an attitude intensity of 1 between *p* and *o* mean to the subjects? In column 6, again, contrary to column 7, a weaker conflict in attitude intensity towards *q* [—, 3, 1] is consistently rated as being more unpleasant than a stronger conflict [—, 3, 2]. This actually goes counter to the psychological principles that demand that the stronger the conflict between friends, the more unpleasant the situation should be experienced.

Column 8 also presents problems. In this column the attitude of *p* to *q* is very weak [1]. All the situations in this column are rated as slightly unpleasant and are very similar to the two situations of column 5: 3, 1, 1 and 1, 1, 1. And again, similar to the situation in column 5 we find a reversal, with respect to the principles, that under these conditions of weak conflict a strong attitude between *p* and *o* is preferred to a weak one. That the attitude intensity between *p* and *q* has no linear effect is evident from columns 6 and 7 where, despite systematic differences in this attitude, the situations have an almost identical range of ratings. But the order is there and the questions it raises are unanswered.

Columns 2, 3, and 4 are equally confusing. The situations in column 3 with mild attitude strength are, as a group, rated as being more pleasant than those in both columns 2 and 4. Column 8 is in a similar position with respect to columns 6 and 7—but what is common to columns 3 and 8? The ordering of the situations within column 3 makes some sense. Given a constant attitude strength of *p* to *q* in this column of +2, then when the attitude of *o* to *q* is slight [+1] the attitude between *p* and *q* becomes irrelevant. With the strengthening of *o*'s attitude to *q* [+3] the attitude between *p* and *o* becomes important and, as to be expected on the basis of the psychological principles, a situation with a strong attitude between *p* and *o* [3, 2, 3] is preferred to a situation with a slight attitude [1, 2, 3]. It is a pity that the situation 2, 2, 3 was not submitted to the subjects for rating; how nice it would have been to find it between the other two.

Column 2 is especially confusing. Four situations are bunched together without any semblance of an order, and then a fifth situation, with stronger positive attitudes than in the other four with a seemingly weak conflict, is rated as quite less pleasant than the others—needless to say in contradiction to the psychological principles. It is possible to construct another order from the two columns just discussed; it is possible to take three situations from each, those where *o*'s attitude towards *q* is +1, and say that these are all slightly pleasant irregardless of the other relations existing, but then how does one account for the striking difference between 3, 2, 3 and 3, 3, 2?

Column 4 is nicely spread out but the order underlying the spread is elusive. Let it be an exercise to the reader to try to establish some order.

One final disturbing point must be raised. The single consistent and significant discriminating point between balance and imbalance in the two earlier experiments was the great sensitivity to change of the balanced situations and the relative insensitivity of the imbalanced situations. Mere visual scanning of Figure 1 shows that this is not the case in the present experiment; if anything, the imbalanced situation here is more sensitive to the experimental changes than is the balanced situation.

From the standpoint of mathematical modelling it would be very nice and neat were the two distributions of scores mirror images of each other. Columns 1 and 5 do mirror each other to a noticeable extent, but there is nothing "mirrory" about the other six columns. Yet, were one to rank order the arrays of ratings in columns 2 and 3, Table 1, and calculate the rank order correlation, one would find it to be —.91. Quite high; quite mirrory. Of what value is this kind of number? Even more, citing this correlation as justification, model builders would rush in to build models that necessarily obscure that which the correlation obscures. Of what value would such a model be?

There is another interesting facet of the data that should be mentioned.

TABLE 4
THE *P-O-Q* SITUATIONS AND THEIR STANDARD DEVIATIONS

SD	+ + −	+ + +	+ − +	+ − −
9	2, 2, 2			
10	1, 1, 1			
11	1, 2, 1	1, 2, 1		
	2, 2, 1			
12	1, 1, 2	2, 1, 2		
	3, 3, 1			
	1, 2, 2			
13	3, 2, 1	3, 2, 2	1, 3, 1	
	3, 2, 2	1, 2, 2		
	2, 3, 1	3, 2, 1		
	2, 3, 2	1, 3, 1		
14	3, 3, 1	1, 1, 2	3, 3, 1	
	3, 1, 2	2, 1, 1		
	1, 3, 1			
	3, 3, 3			
15	3, 3, 2	2, 3, 1	3, 3, 2	3, 1, 1
	3, 2, 3	2, 2, 2	1, 3, 2	
		1, 3, 2		
		1, 1, 1		
		3, 1, 2		
		3, 1, 1		
16	1, 1, 3	2, 2, 1		
17	1, 2, 3	3, 3, 2		3, 3, 3
	2, 3, 3	3, 3, 1		
		2, 1, 3		
		3, 2, 3		
18	1, 3, 2	3, 3, 3		1, 3, 3
	2, 2, 3	1, 1, 3		
		2, 3, 3		
19	3, 1, 3	1, 3, 3	3, 3, 3	
		1, 2, 3		
20	1, 3, 3	3, 1, 3		
21			1, 3, 3	

Note: The data in Table 4 are adapted from "Positive and negative affect as a function of perceived discrepancy in ABX situations" by Price, Harburg, and McLeod (9), and are used by permission of *Human Relations* and the senior author.

Again, a striking but incomprehensible order is found. Table 1 gives the standard deviations for each mean rating. This is a measure of the spread of individual ratings about the mean. Table 4 groups the situations with respect to their standard deviations.

Inspection of Table 4 shows that starting with a given standard deviation magnitude—14 for column 2 and 17 for columns 3, 4, and 5—all situations with larger standard deviations have a strong attitude [3] between *o* and *q* (with two exceptions in column 2) and no situation with a smaller standard deviation has that attitude between *o* and *q*. Why should a strong attitude between *o* and *q* cause so noticeable and consistent an increase in the spread of individual ratings about the mean?

One final comment. In an earlier paper (7) a "strange" pattern of ratings was found to hold in several places, a pattern called "altruism." In quite a few *p-o-x* situations where *p* and *o* had different attitudes towards *x*, subjects "preferred" situations where *p* had a negative attitude towards *x* over those where *p* had a positive attitude. No attempt was made to explain this preference. The data here presented also afford a striking instance of this "altruism." Table 1, column 3, gives the mean rating of six imbalanced situations where *p* dislikes *q*, while *o* likes *q*. In column 2 of the same table we find situations comparable in all other respects except that the attitudes of *p* and *o* towards *q* are reversed. In all six cases the situations where *q* is disliked by *p* are rated as being *less unpleasant* than the comparable situations where *q* is liked by *p*.

E. CONCLUSION

The purpose of this paper was to clarify some issues with respect to research on the *p-o-x* unit and to present some complications that have emerged with research on a *p-o-q* unit. This it has done.

REFERENCES

1. CARTWRIGHT, D., & HARARY, F. Structural Balance: A generalization of Heider's theory. *Psychol. Rev.*, 1956, **63**, 277-293.

2. HEIDER, F. Attitudes and cognitive organization. *J. of Psychol.*, 1946, **21**, 107-112.

3. ———. The Psychology of Interpersonal Relations. New York: Wiley, 1958.

4. HERSHKOWITZ, A. Interpersonal agreement and disagreement on objects and values—A preliminary study. Unpublished Doctoral dissertation, The University of Kansas, Lawrence, Kansas, 1954.

5. JORDAN, N. Behavioral forces that are a function of attitudes and cognitive organization. *Hum. Relat.*, 1953, **6**, 273-278.

6. ———. Perceived discrepancy in attitude intensity between the actors A and B in AB*X* situations and its effect upon affect, or, why don't experimenters utilize all their available data in testing hypotheses? *J. of Psychol.*, 1966, **63**, 299-308.

7. ———. On Cognitive Balance. Research Paper P-178, Institute for Defense Analyses, Arlington, Virginia, 1966.
8. Koffka, K. Principles of Gestalt Psychology. London: Kegan Paul, Trench, Trubner, 1936.
9. Price, K. O., Harburg, E., & McLeod, J. M. Positive and negative affect as a function of perceived discrepancy in *ABX* situations. *Hum. Relat.*, 1965, **18**, 87-100.

Economic and Political Studies Division
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, Virginia 22202

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 23-28.

THE SIGNIFICANCE OF THE EXTINCTION SITUATION*[1]

Department of Psychology, Oklahoma State University

Roy Gladstone

A. Introduction

Kimble (2) reviewed the literature, initiated by Tolman, on latent or nonresponse extinction. He interprets the literature to indicate that extinction can take place in the absence of extinction responses. The negative results of a few studies are interpreted to be a function of the fact that the information supplied the *S*s was ambiguous.

The design of the experiments that showed the phenomenon most clearly involved (*a*) training rats to run to a food reward, (*b*) placing them in the goal box when it was devoid of food, and (*c*) carrying out normal extinction trials. These rats, suitably handled, extinguished faster than control rats. Studies testing the concept with Skinnerian apparatus utilized the secondary reward characteristics (clicks) of the apparatus.

The theory upon which this study is based is that extinction is a function of the significance to *S* of the cues presented during the extinction period. The specific hypothesis is that perception of the exhaustion of the reward reservoir will cause humans to stop responding. Another hypothesis is that no significant difference in extinction will occur when the reward reservoir is hidden *vs.* when it is open to observation but only part of the rewards can actually be obtained.

This study differs from the previous studies in the use of humans rather than rats, in the use of Skinnerian conditioning apparatus without regard to secondary rewarding characteristics, and to a considerable extent in the discussion of theory.

B. Method

1. *Equipment*

A Scientific Prototype Rat Pellet Dispenser dispensed BBs, which were turned in for rewards. The dispenser was modified so that the BBs in the

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 23, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The Research Foundation of the Oklahoma State University has generously underwritten the costs of publication of this article.

magazine could be visible or invisible at the discretion of the *E* and was operated by pulling a knob. The BBs were dispensed on a combination of fixed interval and random-ratio schedule. After one BB fell from the magazine, another would not fall until a 10-second interval had passed followed by a random, small-ratio number of pulls on the knob. The BBs fell into an open, visible box. The dispenser was programmed so that no more than 10 BBs would fall no matter how many were in the magazine.

2. *Subjects*

*S*s were volunteers of college age obtained from classes of the Oklahoma State University. There were 72 *S*s in all, with 24 *S*s in each treatment group. Each group was subdivided by sex, giving 12 *S*s in each cell.

3. *Experimenters*[2]

The experiment was conducted in conjunction with a small class of graduate students in psychology. Six students and the author ran the *S*s.

4. *Procedure*

Sheets were prepared on which volunteers could sign up at times convenient to them, the only restriction being that three consecutive spaces were assigned to a single sex. As the *S*s arrived at the chosen time, *E* assigned them successively to three different treatments. In Treatment 1, *S* could see 10 BBs in the magazine and see them fall, one by one, into an open box. In Treatment 2 there were also 10 BBs in the magazine, but they were hidden from *S*. As in all treatments, however, *S* could see the BBs as they dropped into the box. In Treatment 3, *S* could see 20 BBs in the magazine and see 10 of these drop into the box. Ten remained visible in the magazine but would not fall.

*E*s were instructed to take the *S*s in each group, one at a time, to the experimental room, where the *S* stood in front of the apparatus and was given the following instructions:

> Your task in this experiment will be to operate this machine. Here is how it works. Pull this plunger out and push it in and a small BB will drop into this cup, like this (*E* demonstrates). Later you will be given one stick of gum, some M & M candy, or a cigarette for every BB you have. Do you understand what you are to do?

---

[2] Data were gathered by the following graduate students: Harold Fisher, Jeannette Hoehn, Jean Arthur Kenworthy, Kenneth L. Thompson, Vivian K. Travis, and Richard Washburn. Mr. Thompson suggested the exploratory hypothesis referred to above and Mr. Fisher devised the statistical analysis and carried out the calculations.

E was allowed to repeat the essential instructions for the operation of the apparatus, but questions as to the nature of the experiment were answered pleasantly with the phrase "I am not allowed to tell you any more about the experiment."

E then told the *S*: "You may start now. Please tell me when you are through." (*E* responded to *S*'s question "Can I stop now?" with the phrase "It's up to you.") When the *S* completed his responding, *E* gave the desired reward to *S* in exchange for the number of BBs in the reward cup.

E asked *S* why he stopped responding, and recorded *S*'s answer. *S*s were then asked not to discuss the experiment with anyone else.

*E*s practiced the procedure on one another before running any *S*s. The number of pulls in the extinction period were recorded automatically. In a few cases, *S*s in the 10-invisible and 20-visible groups were stopped after they had pulled the knob more than 500 times in the extinction period and showed no signs of stopping.

C. Results

Some *S*s stopped after a few BBs had fallen. These were eliminated from consideration, resulting in uneven *N*s in the cells. The *N* in each cell was reduced to 12 through the use of a table of random numbers. One *S* was removed in each of three cells and two *S*s from one cell for that purpose.

The mean number of pulls in each cell appears in Table 1. The results of the analysis of variance testing the differences between those means appears in Table 2. The mean number of pulls in the 10-visible condition was significantly smaller than in either of the other conditions; the 10-invisible condition produced significantly fewer pulls than the 20-visible condition, and sex was not a significant variable at the .05 level.

TABLE 1
Mean Number of Pulls

Condition	Male	Female	Combined
1 (10 visible)	32.17	36.50	34.33
2 (10 invisible)	224.33	118.42	171.38
3 (20 visible)	312.00	223.50	267.75

Some of the most interesting data cannot be neatly summarized. In the 10-visible condition, two *S*s yielded no extinction pulls at all, three yielded 1, and two yielded 2. It is likely that some of the latter would have yielded a zero extinction score except for the fact that the *S* tended to skid past the stopping point when responding quickly. By contrast, the two lowest extinction

TABLE 2

TREATMENT COMPARISONS

Source of variation	Degrees of freedom	Sum of squares	Mean square	*F* ratio
Mean	1	1,793,302.24		
Conditions	2	660,415.20	330,207.60	17.821**
C_1 (1 *vs.* 2)	1	225,365.02	225,365.02	12.162**
C_2 (1 *vs.* 3)	1	653,800.08	653,800.08	35.284**
C_3 (2 *vs.* 3)	1	116,457.69	116,457.69	6.285*
Sex	1	72,263.35	72,263.35	3.899
Conditions × Sex	2	42,152.86	21,076.428	1.137
Experimental error	66	1,222,953.25	18,529.56	

* Significant at the .025 level.

** Significant at the .0005 level.

scores in either of the other conditions were 9 and 10; the next lowest score was 25.

The answer to the question "Why did you stop?" varied between the conditions. In the 10-visible condition some *S*s said, "Because there were no more BBs," and some indicated they had not stopped when all the BBs had fallen because they suspected some trick. In contrast, many in the other conditions said they stopped because no more BBs were coming out or because they were tired.

D. DISCUSSION

The results of this experiment can be explained best by Tolmanian theory. They can be explained indirectly by a few other concepts designed to explain extinction phenomena, and cannot be explained at all by other pertinent theories. Since different extinction phenomena emerged from the same set of learning conditions, the results cannot be explained by pointing to differences in reinforcement schedules. Since many of the *S*s in the 10-visible condition yielded no or very few extinction Rs, and since there were differences in means between the groups, the concepts of reactive and conditioned inhibition cannot be used to explain the results except via the generalization of a lack of response to emptiness. The latter concept would require a history of *S*'s making long series of Rs to emptiness, becoming tired, and being conditioned not to respond to the stimulus of emptiness by the reinforcement inherent in doing nothing when tired or by the association of doing nothing with the stimulus of emptiness, and then generalizing that R to this new situation. Whether or not this picture is valid remains to be established by future experimentation.

The concept of secondary reinforcement seems difficult to accept in explaining the results. The secondary reinforcing properties of the room, the *E*, and the apparatus should be massive compared to the difference made by the drop of the last BB. Why should several *S*s stop abruptly at that point? Why should the BB in the box not have the same secondary reinforcing properties as when it is in the magazine? There seems to be no good answer to these questions within the concept of secondary reinforcement.

The concept of frustration will not explain the results. *S* is no more frustrated after the last visible BB drops than after the last invisible BB; yet he stops in one situation but not in the other.

The discrimination hypothesis as stated with a behavioristic cast by Bitterman, Fedderson, and Tyler (1) also seems hard to accept when used to explain the results of this experiment. They state the discrimination hypothesis as follows: ". . . the rate of extinction is inversely related to the similarity between conditions of training and extinction." The visible drop of the last BB hardly seems to make the situations so quantitatively dissimilar as to make as much difference as it does. However, the disconfirmation of the quantitative aspect of the stated hypothesis is not as clear as it might be and another experiment is being planned that is designed to deal with the matter.

Hypotheses couched in terms of the proprioceptive consequences of reward are not satisfactory, since some *S*s stop responding abruptly after rewards stop and the proprioceptive consequences of lack of reward cannot have made themselves apparent. In addition, the proprioceptive consequences concept would lead to a prediction of the same mean number of extinction pulls in all three conditions, a situation that did not occur.

Finally, the original hypothesis that the perception of emptiness will cause human *S*s to stop responding is confirmed by the significance of the difference between means, although not by every datum considered individually.

E. EPILOGUE

This experiment has shown that the projection of a particular cognitive response upon other human beings leads to better predictions of behavior with this population than predictions based upon behavioristic principles (although it is possible to explain the learning of that cognitive *R* by behavioristic principles). It seems intuitively obvious that projection might be used to predict the consequences of many situations involving school children at all levels. It is also true, however, that many such projections will lead to very poor predictions. The problem is to discover the characteristics of the situations in which such projections may be made with a reasonable degree of confidence. One approach to the problem is to try to discover those learned responses that are universal, or at least almost universal, within given cultures. Piaget has led the way. The exploration of the characteristics of responses to emptiness is in the tradition devised by him.

REFERENCES

1. BITTERMAN, M. E., FEDDERSON, W. E., & TYLER, D. W. Secondary reinforcement and the discrimination hypothesis. *Amer. J. Psychol.*, 1953, **66**, 456-464.
2. KIMBLE, G. A. Hilgard and Marquis' Conditioning and Learning. New York: Appleton-Century-Crofts, 1961. Pp. 320-323.

Department of Psychology
Oklahoma State University
Stillwater, Oklahoma 74075

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 29-32.

THE PAIRED ASSOCIATE TASK AS A PREDICTOR OF FOREIGN LANGUAGE FLUENCY*

University of New Mexico and San Jose State College

ARTHUR C. VICORY[1] AND JAMES J. ASHER

A. INTRODUCTION

The primary purpose of this research has been to determine to what extent competence in a foreign language can be predicted from measures of paired associate learning.

Although the paired associate task has been widely applied by psychologists and psycholinguists to foreign language learning and prediction (1, 2, 3, 4, 5, 6), most linguists have objected to the findings as not representing the terminal behavior goal of training the student to "think" in the language. Asher (1) has asserted that the paired associate model adequately represents foreign language fluency empirically, although it may not be related rationally to "popular" linguistic methods of teaching foreign languages. The question of a rational relation becomes less relevant if it is determined empirically that paired associate learning predicts language fluency, because the *predictive validity* of the paired associate model may be of pragmatic value in further laboratory and classroom research. This study attempted to obtain empirical evidence for the paired associate hypothesis.

B. METHOD

1. *Subjects*

Thirty-six male United States Army personnel at the Presidio of Monterey Army Language School served as *S*s in an 11-month intensive language training course.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 23, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This article is based on a thesis submitted in partial fulfillment of the requirements for the M.A. degree in psychology while the senior author was at San Jose State College, California. The author is indebted to James J. Asher for his patient guidance in all phases of this research. Sincere thanks are given to William N. McBain and Jay T. Rusmore for their helpful suggestions. Special acknowledgment is made to Bela Banathy, Director of East Europe-Middle East Division at the Presidio of Monterey Army Language School, and to several members of his staff for their helpful cooperation in providing the language student samples and facilities necessary for this investigation.

2. *Procedure*

The 36 *S*s were screened from a larger sample on the basis of scores made on general intelligence and language aptitude tests. Each *S* was assigned to learn either Polish (two groups of 9 *S*s each), Hungarian (9 *S*s), or Bulgarian (9 *S*s). The four groups were matched on age, formal education, intelligence, previous foreign language experience, and language aptitude. Another series of intelligence and language aptitude tests were given to the *S*s during two days immediately prior to language training.[2] Included in this test battery were two paired associate language predictors, which were group administered in the following order: (*a*) Polish Aural Paired Associate (Aural PA) and (*b*) Polish Visual Paired Associate (Visual PA). The Aural and Visual PA tests each consisted of a different list of 10 paired items, an English stimulus word (e.g., other) and a Polish response word (e.g., *inny*). Stimulus and response exposure time were equated for the two sense modalities by employing an anticipation procedure, where the stimulus was given for two seconds followed by a seven-second *S* response interval (stimulus not presented); then response confirmation was given for two seconds. Interitem interval was seven seconds. *S*s observed or listened without responding for two familiarization trials, followed by 15 responding trials. Written responses were recorded in booklets during Visual PA learning, while oral responses were tape-recorded during taped Aural PA learning. Response scoring was done by the experimenter.

All *S*s received intensive eight-hour-a-day study five days a week for 11 months. Audio-lingual skills were emphasized throughout the course. Each language was taught by three to five instructors, with each instructor teaching only one language. Three criterion measures were collected after one, two, and six months of language learning: (*a*) Daily grades were assigned to each student by their instructors, and cumulative grades (CG) were computed monthly. (*b*) At the end of each month the instructors made paired-comparison (PC) ratings for each student by comparing them with one another on language achievement. (*c*) Following the PC ratings, each of the Polish groups was given the Polish Language Achievement Test (PAT). A similar test was not available at the time of this study for the Hungarian and Bulgarian groups.

C. RESULTS AND DISCUSSION

The main findings in support of the paired associate hypothesis may be summarized as follows:[3] (*a*) Only the Aural PA predictor showed a consistent pattern of significant validity coefficients with the criterion measures collected throughout the six months of language study, and only with the Polish group shown in Table 1. (*b*) The highest validities occurred with the instructors' PC ratings. (*c*) All three criteria collected after six months were significantly correlated with Aural PA learning. (*d*) The Visual and Aural PA predictors were not significantly correlated with one another for any language group.

---

[2] Only selected tests will be discussed, since this article is primarily concerned with the paired associate predictor.

[3] The analysis of much of the data was made possible through the facilities of Western Data Processing Center at the University of California, Los Angeles.

TABLE 1
VALIDITY COEFFICIENTS FOR POLISH PAIRED ASSOCIATE PREDICTORS (AURAL OR VISUAL PA) WITH THE PAIRED COMPARISONS (PC), CUMULATIVE GRADES (CG), AND POLISH ACHIEVEMENT TEST (PAT) CRITERIA[a]

	Criteria						
	PC			CG[c]		PAT[c]	
Predictors	1st	2nd	6th	1st	6th	1st	6th
Aural PA ($N = 7$)[b]	1.00**	1.00**	.89**	.45	.75*	.42	.87**
Visual PA ($N = 7$)[b]	.17	.16	.54	.62*	.69*	.48	.47

* Significant at .05 level.

** Significant at .005 level.

[a] The results are shown for only one Polish group, since the other Polish group received additional training during the first month that will not be evaluated.

[b] Sample size was reduced from 9 to 7 because two student dropouts occurred during the second month of training.

[c] The CG and PAT criterion measures were not available for the second month.

Although the results of Table 1 appear to be highly significant, a cautious interpretation should be made, since the standard error of small samples may reduce reliability. This possibility is minimized to some extent, since the paired associate and criteria reliability estimates were high. Paired associate reliability averaged .995. The PC ratings were the most reliable of the criteria (average .980 by Fisher's *z* transformation).

None of the other predictors correlated highly with Polish except the Polish Language Aptitude Test, which is a "specific language aptitude" predictor employing natural language material.

D. CONCLUSIONS

Limited support was found for the paired associate task as a predictor of foreign language fluency. The evidence tends to support the following inferences: (*a*) the verbal response units of the paired associate task should be selected from the particular natural language for which fluency is predicted, and (*b*) the Aural PA task provides the best prediction when the above condition is satisfied. These inferences must be considered speculative until further research not only cross-validates the findings with Polish, but demonstrates similar results with other natural languages.

REFERENCES

1. ASHER, J. J. The paired-associate task: Adequate or inadequate in psycholinguistic research? Paper read at the Western Psychological Association Meeting, San Francisco, California, April 1962.
2. ———. Sensory interrelationships in the automated teaching of foreign languages. *Percept. & Motor Skills,* 1962, **14**, 38.
3. ———. Vision and audition in language learning. *Percept. & Motor Skills,* 1964, **19**, 255-300 (Monog. Suppl. 1-V19).
4. CARROLL, J. B. A factor analysis of two foreign language aptitude batteries. *J. Gen. Psychol.,* 1958, **59**, 3-19.
5. PIMSLEUR, P., & BONKOWSKI, R. J. Transfer of verbal material across sense modalities. *J. Educ. Psychol.,* 1961, **52**, 104-107.
6. VICORY, A. C., & ASHER, J. J. A simplified associative-frequency measure of meaningfulness in verbal learning. *J. Verb. Learn. & Verb. Behav.,* 1966, in press.

HumRRo
Division No. 5 (Air Defense)
P. O. Box 6021
Fort Bliss, Texas 79916

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 33-40.

MATCHING PATIENTS AND THERAPISTS: A METHODOLOGICAL APPROACH*

Department of Psychiatry, University of Utah College of Medicine

ROGER B. ALLISON, JR., IJA N. KORNER, AND MAX D. ZWANZIGER

A. INTRODUCTION

Luborsky and Strupp (2) consider "the match between patient and therapist" to be among the more promising areas reported in their summary of the 1961 Conference on Research in Psychotherapy. This area includes those characteristics (e.g., social class, expectancies, values, etc.) of the patient and the therapist that hopefully lead to cooperative and productive psychotherapy. Our present-day procedures for clinical evaluations—interviews, social history, psychological tests—permit an extensive description of characteristics of the patient. How pertinent these characteristics are with respect to the problem of matching patient and therapist has not been demonstrated. Complete as its description of the patients may be, the typical research study in psychotherapy provides little information about the characteristics of the therapists. Sundland and Barker (3), for example, report that the usual description of therapists lists age, sex, years of experience, professional training, and "school" of psychotherapy. They further conclude that "none of the studies has resulted in a comprehensive measure of the therapist variable in the therapy situation" (3, p. 201). It would seem that the general impression that the relationship between therapist and patient dwarfs profession, experience, "school," dynamics, and diagnosis would encourage research in psychotherapy to move toward specifying the characteristics of both patients and therapists, as well as the idiosyncracies of combination that facilitate the achievement of therapeutic goals.

In the present paper, the authors are suggesting and illustrating a methodological approach for organizing therapists into subgroups based upon their similarities and differences. The method is easily extended to form subgroups of patients, or subgroups of *both* patients and therapists.

The method is essentially one of inverted factor analysis—inverted because the correlation matrix consists of the interrelationships among the subjects based upon the scores or ratings on many variables. This is in contrast

---

* Received in the Editorial Office, Provincetown, Massachusetts, on June 3, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

to the traditional factor analytic model, which studies the interrelationships among the variables rather than the subjects. (Most standard textbooks on factor analysis discuss this procedure, and computer programs for both the factoring and rotational stages are readily available. The difficulties of the mathematics of factor analysis, therefore, no longer justify an investigator's avoidance of the technique.)

The inverted factor analysis begins with a score matrix, which is simply the scores earned on a large number of variables for each of the subjects under study. These scores should have roughly comparable units of measurements, which may warrant normalizing the frequency distributions of scores for each variable and converting the normalized scores to standard (or normal deviate) scores. The subjects are then intercorrelated with each other by conventional procedures. Since the standard error of the resulting correlations will be a function of the number of variables used, the best estimates of the relationships between subjects will come from having a large number of variables. The factoring and rotational stages will produce the amount that each subject has for each of the derived factors. These factors are hypothetical and are then defined or explained on the bases of the *similarities* of the subjects within a factor and the *differences* of the subjects *between* factors.

Inasmuch as the information obtained from factor analysis depends upon the information available in the score matrix—that is, the researcher cannot expect to find information in the factor analytic solution that was not present in the score matrix—attention needs to be focused upon the subjects to be studied and the specific variables that will provide the bases for similarities and differences among the subjects. This is the heart of factor analytic design and the researcher has two options. (*a*) He can use factor analysis to test hypotheses about the nature of the organization of the subjects, in which case he will use three or more subjects who on an *a priori* basis are alike for each of the hypothesized factors. For example, if it were hypothesized that therapists would be organized on the basis of their respective "schools," then the subjects might include three or more therapists from each of several "schools" (Freudian, Rogerian, etc.). The variables should then be selected that bear upon fundamental or related aspects of the so-called schools. (*b*) The researcher can use factor analysis to create hypotheses, which is highly desirable whenever little knowledge is available for a particular domain. In this case, the subjects may simply be therapists and the variables may be selected to provide information pertinent to the practice of psychotherapy or characteristic of people in general. The factor analysis will usually lead to the formation of a number of subgroups of therapists, and subsequent interpretation of the identity of the subgroups will then constitute hypotheses for further investigation.

The remainder of this paper illustrates and discusses the use of the second option—namely, the development of hypotheses about the nature of the organization of therapists. Patients could be studied in a similar manner. Subgroups could also be formed by including both patients and therapists in the factor analysis.

B. ILLUSTRATIONS OF THE METHOD

The following two examples represent attempts to form subgroups among psychotherapists. The first example is based upon the ratings of six therapists and one patient during a unique psychotherapeutic situation, which has been described by Korner *et al.* (1). Briefly, one patient was seen for 13 consecutive hours in psychotherapy by each of seven psychotherapists, who rotated in two one-hour sessions with one hour rest in between hours. One therapist was dropped from the analysis because he served for only one hour. After each hour, the therapist completed a 23-item checklist pertaining to the progress of the hour, an evaluation of the patient's behavior, and an evaluation of the therapist's behavior. An independent set of six observers (each of whom had training and experience as psychotherapists) also rotated assignments observing each hour in an adjacent room. The observers completed the same checklist. Thus, each therapist and observer had completed two checklists, or a total of 46 items pertaining to the practice of psychotherapy. From a review of the films and tape recordings made of the therapy sessions, general consensus was reached that the patient failed to show appreciable reduction or change in her resistances. Subsequent testing and follow-up interviews during the next 20 months further confirmed the absence of change in the patient's pathology and defenses. She was therefore considered to be an experimental constant (a test, if you like) to which 12 subjects provided 46 responses.

The researchers then sought to organize the 12 subjects into subgroups on the bases of the similarities and differences of their ratings to the 46 checklist items. The 12 subjects were intercorrelated and the matrix of intercorrelations was factor analyzed and rotated to simple structure. The rotated factors were relatively uncorrelated with each other. The results appear in Table 1.

Essentially, the factor analysis based upon the checklist items implied that the 12 therapists and observers could be organized into three basic groups of people. Similarities are represented *within* a factor. That is, subjects 1, 2, 7, 8, and 11 are similar to each other. Differences are represented *between* factors. That is, the subjects listed under Factor I are different from those listed under Factor II, and so forth. The numbers listed in the "loading" column provide a measure of how much of the hypothetical factor each subject has. In this case values less than .50 are not reported because they are less important than values above .50 in terms of identifying the factors, and are perhaps not significantly different from zero.

This is as far as the mathematics of factor analysis can take us. What we want to know is what these three factors mean. In what way are the people alike who "have" Factor I and in what way are they different from those who "have" Factor II, and so on?

TABLE 1
FACTOR ANALYSIS BASED UPON CHECKLIST

Factor and subject	Loading
Factor I:	
1. Th Sr Psychologist AB	.52
2. Th Sr Psychologist A	.57
7. Ob Jr Psychiatrist B	.82
8. Ob Sr Psychiatrist AB	.63
11. Ob Jr Psychiatrist A	.64
Factor II:	
3. Th Sr Psychiatrist B	.81
4. Th Sr Psychiatrist AB	.74
Factor III:	
5. Th Jr Psychiatrist A	.56
6. Th Jr Psychiatrist AB	.76

Note: Th = therapist, Ob = observer, Sr = senior, and Jr = junior.

Table 1 also lists some of the obvious characteristics of the subjects: whether they were therapists or observers, whether they were senior or junior (a relative and arbitrary measure of experience), and whether they were psychiatrists or psychologists. The Betz-Whitehorn classification of therapists (4) was added for the 11 male subjects. Subjects with "A" ratings reportedly work well with schizophrenics, while those with "B" ratings do not. An "AB" rating is midrange in the A to B continuum. Note that the Betz-Whitehorn ratings appear to be randomly distributed across the three factors and, therefore, do not help in the identification of the factors.

If we exclude the observers for a moment, the identification of the factors is quite distinct. Factor I consists of two senior psychologists; Factor II, of two senior psychiatrists; and Factor III, of two junior psychiatrists. The authors concluded that the major bases for making the ratings came from the training and experience of the individual who served as a therapist. Not as clear are the ratings produced by the observers. The three psychiatrists who served as observers behaved much like the experienced psychologists who served as therapists. The authors felt that this might represent some increase in objectivity, which could result from not being directly involved in the therapeutic transaction. Moreover, since the psychologist's image of himself includes objectivity, this would account for the high loadings of the psychiatrist-observers on the psychologist factor. However, the observer may note that three subjects who served as observers are missing from the table. These three subjects were all psychologists and did not have loadings of .50 or more on any factor, including Factor I which appears to be the psychologist factor.

However, these three psychologists did have moderate loadings on Factor I, supporting the interpretation that Factor I represents a psychologist factor. They also have moderate loadings on the other two factors, suggesting that they function in part as psychologists, and in part as psychiatrists. The main conclusion the authors reached from this factor analysis is that ratings of psychotherapy are influenced by both training and experience.

The second example is based upon an analysis of the 12 subjects in terms of characteristics that are not specific to psychotherapy. This time the subjects were interrelated on the basis of 49 scales from the Strong Vocational Interest Blank. Subject 12, who was female, was dropped from this analysis for lack of comparable scores brought about by separate scales for men and women.

The results of the factor analysis based upon the Strong scales are presented in Table 2. Once again the authors found three factors, none of which was defined by the Betz-Whitehorn system. Note from the table that the authors are dealing with a vastly more complicated organization than resulted from the checklist ratings. None of the characteristics listed for the subjects provides a key to understanding the factors. Since the factors were determined from the Strong scales, the authors returned to a careful study of each scale to clarify the factors.

Before stating the conclusions about the nature of the three factors, it might be helpful to review some of the relationships that were found between the Strong scales and the factors. For example, Factor IV was highly correlated with the Lawyer scale and negatively related to the Junior Accountant scale. This factor was also related to a lesser degree to the Architect, Author-Journalist, and Occupational Level scales, and negatively to the Office Worker scale.

Factor V was a bipolar factor highly correlated with the Life Insurance Salesman scale. Individuals who had high positive loadings on this factor had low scores on the Salesman scale, while those who had low factor loadings were in the middle range of the Salesman scale, and the individual with a high negative loading on this factor scored highest on the scale.

Factor VI was negatively related to the Architect and Author-Journalist scales. This factor correlated — .40 with Factor IV, as would be suspected from the dual relationship of the Architect and Author-Journalist scales.

The authors believe that Factors IV and VI—which are negatively correlated with each other—reflect the drive level of the individual. The individual thinks of himself as an ambitious and humanistic man with broad interests and intellectual cravings. What differentiates Factor IV from Factor VI is the yearning for prestige. While individuals with high loadings on Factor IV seek prestige in public recognition and demonstration, individuals high in Factor VI prefer the nonpublic recognition of peer or more general insider groups.

TABLE 2
FACTOR ANALYSIS BASED UPON STRONG

Factor and subject	Loading
Factor IV:	
2. Th Sr Psychologist A	.70
4. Th Sr Psychiatrist AB	.66
5. Th Jr Psychiatrist A	.81
8. Ob Sr Psychiatrist AB	.80
11. Ob Jr Psychiatrist A	.83
6. Th Jr Psychiatrist AB	.57 (Also loads on V)
10. Ob Jr Psychologist A	.69 (Also loads on V)
Factor V:	
7. Ob Jr Psychiatrist B	.72
6. Th Jr Psychiatrist AB	.53 (Also loads on IV)
9. Ob Sr Psychologist B	.53 (Also loads on VI)
10. Ob Jr Psychologist A	—.63 (Also loads on IV)
Factor VI:	
1. Th Sr Psychologist AB	.72
3. Th Sr Psychiatrist B	.74
9. Ob Sr Psychologist B	.52 (Also loads on V)

Note: Th = therapist, Ob = observer, Sr = senior, and Jr = junior.

Factor V might well be related to the individual's need to relate to people and the manner in which these needs are to be met. Individuals high on this factor (and, therefore, low on the Life Insurance Salesman scale) need immediate acceptance in a warm, friendly environment. Individuals with high negative loadings on this factor prefer to "sell" themselves in a slightly hostile environment. Individuals with low loadings are probably less concerned than are individuals with high loadings about the positive and negative attitudes of their environment.

In summary, the factor structure that resulted from the Strong data resembled a structure that was not at all confined to psychotherapists. It is likely that a similar organization might be found among lay people.

C. IMPLICATIONS AND DISCUSSION

The above factor analyses illustrate a method by which therapists could be organized into subgroups and could be studied further in terms of therapeutic efficacy with different types of patients. Although these studies are not definitive with respect to the specific factors, they suggest that the organization of therapists may pivot around two basic issues. One relates to the practice of psychotherapy in which concepts of movement, dominance, resistance, transference, etc., provide the basis for grouping psychotherapists. In this case, training and experience variables would naturally play important roles. Since training and experience obviously change with time, the factor analytic organization would not necessarily be stable and may, therefore, have limited value for matching therapists and patients.

The second issue pertains to the organization that results when the characteristics under study reflected properties of people in general—that is, the characteristics were not unique to psychotherapists. Such was the case with the factor analysis based upon the Strong Vocational Interest Blank. Since measurement instruments of stature comparable to the Strong instrument usually report satisfactory test-retest reliability coefficients, a factor structure based upon this type of rating would probably be more stable than one utilizing psychotherapeutic concepts. If so, and if the factor structure were demonstrated to be related to therapeutic efficacy with different types of patients, training institutions would be in the advantageous position of having a ready-made selection program for potential therapists in terms of community needs. This would also suggest that training and experience act as catalysts in developing interpersonal relationships of therapeutic significance.

The first example was obtained from multiple ratings of several therapists for one patient. The analog for determining the organization of patients based upon the practice of psychotherapy might well be found within a group psychotherapy session. Here, many patients could provide multiple ratings for one therapist who, in this case, would be considered to be an experimental constant. The analog for the second example requires only the adoption of group or individual testing procedures.

The most convenient approach to investigating the organization of *both* patients and therapists would be accomplished with conventional paper-and-pencil tests. For example, the Strong Vocational Interest Blank, the Thurstone Interest Schedule, the Edwards Personal Preference Schedule, the Allport-Vernon Scale of Values could each (or collectively) be used to provide similarities and differences among patients and therapists. Once the subgroups have been formed, the problem of successful matching (mismatching or complementing) can be investigated.

In addition to the fascinating questions related to the problem of matching, the questions of stability and change come to the fore. Do patients change toward a stable therapist? Do both patients and therapists change in the direction of a compromise? Does group psychotherapy result in a more homogeneous or heterogeneous collection of people? Is the group therapist in the center of the system, the periphery, or outside? These are but a few of the research studies that could originate from the proposed methodology.

D. Summary

The method of inverted factor analysis was presented as a technique for organizing therapists and patients into subgroups. This approach would allow research into the problem of matching patients and therapists. Two examples were presented to illustrate the method: one typified the practice of psychotherapy; the other was characteristic of people in general. A discussion of the consequences of the approach ranged from the selection of therapists to the nature of changes occurring in group psychotherapy.

References

1. Korner, I. N., Allison, R. B., Jr., Beier, E., Broberg, C., & Zwanziger, M. The patient as a constant in psychotherapy. *J. Clin. Psychol.,* 1964, **20**, 403-406.
2. Luborsky, L., & Strupp, H. H. Research problems in psychotherapy: A three-year follow-up. In H. H. Strupp & L. Luborsky (Eds.), *Research in Psychotherapy.* Washington, D. C.: Amer. Psychol. Assoc.
3. Sundland, D. M., & Barker, E. N. The orientations of psychotherapists. *J. Consult. Psychol.,* 1962, **26**, 201-212.
4. Whitehorn, J. C., & Betz, B. J. Further studies of the doctor as a crucial variable in the outcome of treatment with schizophrenic patients. *Amer. J. Psychiat.,* 1960, **117**, 215-223.

Department of Psychiatry
University of Utah College of Medicine
50 North Medical Drive
Salt Lake City, Utah 84112

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 41-48.

REALITY AND NONREALITY IN STUDENT EXPECTATIONS OF COLLEGE*[1]

Princeton University

Lawrence A. Pervin

A. Introduction

Recent research in psychology has indicated that the expectations an individual brings to a situation significantly influence how he experiences and copes with that situation (2, 3). Concerning the college situation, Thistlethwaite (5) has found that many students who drop out of a field do so because their expectations of certain career fields prove to be incompatible with reality. While there are thus suggestions as to the importance of the expectations variable, the relationship of expectations to satisfactions with and management of new situations, such as the college environment, remains in need of further investigation.

The research reported here is part of a larger investigation of student performance and health as a function of individual-environment interaction. This larger study is in the process of following a class of Princeton undergraduates through four years of college, while assessing the nature of the transactions between individual students and the different parts of the college. The present report deals with one aspect of these transactions: the nature of the expectations that students bring to college and the extent to which these expectations are confirmed or disconfirmed.

B. Method

1. *Subjects*

The main group of subjects consisted of 670 entering freshmen from the Princeton class of 1968. Reports on the Princeton college environment or Princeton press were obtained from 190 undergraduates (50 in each of the freshman, sophomore, and junior classes, and 40 seniors), who were randomly chosen from the larger undergraduate body.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on June 9, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by Grant MH 08321 of the National Institute of Mental Health. Willard Dalrymple, Director of the Princeton Health Services, is co-investigator on the project.

2. *Materials and Procedure*

The College Characteristics Index (CCI) was used as a measure of student expectations about the Princeton environment, and as a measure of the press of the Princeton environment as perceived by students at the college. The CCI is a measure of 30 kinds of press relating to activities, policies, procedures, attitudes, and impressions that characterize an undergraduate college setting (4). The 30 press dimensions have been factor-analyzed into 11 factors. The Past Performance Questionnaire (PPQ) was used as a measure of students' expectations about their own performance at college. This questionnaire inquires into past eating, sleeping, and work difficulties, etc., and also where students expect to stand in their class and how much difficulty they expect to have at college.

In May, 1964, the CCI was administered to 190 Princeton undergraduates who responded according to their perceptions of the Princeton environment. In June, 1964, all accepted applicants ($N = 800$) to Princeton were mailed the CCI with instructions to fill out the questionnaire according to their expectations of Princeton. An accompanying letter indicated that the questionnaire was part of a larger study being carried out by the University Health Services. Students were assured that their responses would in no way affect any administrative decisions and that the information was confidential. The fact that the CCI was completed on a voluntary basis gave further support to the research aspects of the project.

In September, 1964, students were administered the PPQ as they came through the University Health Services during freshman orientation week. The assurances of confidentiality that accompanied the CCI also accompanied the PPQ. Voluntary responses were obtained from 719 students.

C. RESULTS

Means on the 30 CCI scales were computed for the expectations of the entering freshmen (1968e). In analyzing the scale means for the students already at Princeton, it was found that perceptions of the Princeton press changed in a fairly regular manner from the freshman through the senior years. Therefore, means for the four classes (1964-1967) and means for the freshman class alone (1967) were used as measures of the Princeton press to which the expectations could be related.

It is possible to compare the 1968 expectations with the 1967 and 1964-1967 perceptions of the Princeton press in two ways. The first comparison is based on a rank-order correlation and indicates how realistic the expectations are in terms of the relative weighting of different presses. The rank-order correla-

TABLE 1
RAW MEANS ON CCI SCALES FOR ENTERING STUDENTS (1968e) AND STUDENTS AT PRINCETON (1964-1967, 1967)

Scale	1968e ($N = 670$)	1964-1967 ($N = 190$)	1967 ($N = 50$)
1. Abasement	2.52	1.81***	1.54***
2. Achievement	8.82	6.62***	6.88***
3. Adaptability	4.46	3.47***	3.68***
4. Affiliation	7.28	5.52***	5.76***
5. Aggression	4.00	5.95***	6.02***
6. Change	7.51	7.63	7.78
7. Conjunctivity	7.92	7.72	8.30
8. Counteraction	6.41	5.47***	5.54***
9. Deference	4.01	3.55**	3.68
10. Dominance	5.54	3.70***	4.18***
11. Ego achievement	8.22	6.97***	7.54*
12. Emotionality	7.26	5.76***	6.84
13. Energy	7.94	6.50***	7.10**
14. Exhibitionism	7.53	5.73***	6.50***
15. Fantasied achievement	6.47	6.29	6.34
16. Harm avoidance	4.82	4.20***	4.10***
17. Humanities	8.86	8.15***	8.02***
18. Impulsiveness	6.53	6.39	6.88
19. Narcissism	5.43	2.95***	3.20***
20. Nurturance	6.03	4.70***	5.28**
21. Objectivity	8.03	8.36**	8.34
22. Order	5.94	4.62***	5.00***
23. Play	6.44	5.64**	5.98
24. Practicalness	5.33	4.44***	5.30
25. Reflectiveness	9.01	7.34***	7.50***
26. Science	9.04	7.98***	8.18***
27. Sensuality	6.64	6.07***	5.92*
28. Sexuality	5.25	5.07	5.54
29. Supplication	5.42	4.69***	4.98
30. Understanding	8.88	7.76***	7.94***

* $p < .05$, two-tailed test.
** $p < .01$, two-tailed test.
*** $p < .001$, two-tailed test.

Note: 1964-1967 means were tested against 1968e means, and 1967 means were tested against 1968e means.

tion of the 1968e scale scores with the 1964-1967 scale scores and the 1967 scale scores was .74 ($p < .001$).[2] The second comparison is based on means for the 30 CCI scales. These data are presented in Table 1 and indicate that there are many significant differences between expectations and perceptions of the Princeton environment. A comparison of the 1968e group with the 1964-1967 group shows that there are 25 significant differences on the scales, and all of the differences on the 11 factors are significant beyond the .001 level. The differences between the 1968e group and the 1967 group are not quite as large, but there are 19 significant scale differences and 10 of the factors had differences significant at or beyond the .05 level of confidence.

---

[2] The raw scale score means were converted to standard scores for computing the rank-order correlation.

One section of the PPQ dealt with students' expectations of their own performance at Princeton. Students indicated in which decile of class standing they expected to stand upon graduation. They also indicated the minimum decile with which they would be satisfied. Table 2 presents the percentage of students indicating each of the deciles for these two questions. Since 86 per cent indicated that they expect to stand in the top half of the class and 76 per cent indicated that the top half of the class was the minimum acceptable level, a minimum of 36 per cent of the students will fall below their expected level and a minimum of 26 per cent will fall below their minimum acceptable level unless some changes in these levels are made. Since some of the students who expect to fall in the bottom half of the class will actually fall in the top half, and some of the students who expect to stand high in the top half will actually stand low in the top half, the number of students who will fall below the expected and minimum acceptable levels is considerable.

D. Discussion

The data on the expected and actual press of the Princeton environment indicate that both reality and nonreality are present in student expectations. The expectations were quite accurate in terms of the relative distribution of presses at Princeton, but quite inaccurate in terms of the amount of each press. The latter supports the view of Webb (6) that students have relatively vague and unrealistic conceptions of what to expect at college.

Further evidence of the inaccuracy of student expectations comes from

TABLE 2
Expected and Minimum Acceptable Levels of Class Standing upon Graduation Reported by Entering Freshman ($N = 719$)

Decile	Per cent reporting expected level	Per cent reporting minimum acceptable level
Top		
1	8	4
2	16	11
3	24	18
4	18	16
5	20	27
6	7	12
7	4	6
8	2	3
9	1	2
10	0	1
Bottom		

responses to the PPQ. When the students first arrived at Princeton, 69 per cent reported that they had fairly clear or very clear expectations about academic aspects of Princeton, and 56 per cent reported fairly clear or very clear expectations concerning nonacademic aspects of Princeton. One year later these students were asked how clear their expectations were at the time of arrival at Princeton. At this time 56 per cent indicated that their expectations concerning academic aspects had been fairly clear or very clear, and 27 per cent indicated that their expectations concerning nonacademic aspects had been fairly clear or very clear. The change for academic aspects is significant at the .01 level, and that for nonacademic aspects at the .001 level. Similarly, at the end of the freshman year 33 per cent of the students reported that Princeton was somewhat or very different from what they had expected in academic aspects, and 54 per cent reported such differences for nonacademic aspects.

Where were the expectations particularly inaccurate? A comparison of the scale ranks for expected press and perceived press indicates the following: (*a*) Students tended to find *comparatively* less press for intellectual and abstract understanding and less press for group spirit, togetherness, and social welfare than they had expected. (*b*) Students tended to find *comparatively* more press for change, fantasied achievement, objectivity, open-mindedness, and heterosexual interests than they had expected. However, these statements reflect the relative strength of the various presses. If absolute strength is considered, then the entering students expected to find more of almost every press than was perceived by students already at Princeton. Compared to the means for perceived press, those for expected press were particularly high on the following factors: (*a*) Social Form—group activities, warm and friendly relationships, social welfare, social skills; (*b*) Academic Organization—high organization, structure, order; (*c*) Self-Expression—development of leadership potential and self-assurance; (*d*) Group Life—togetherness, devotion to the welfare of others; (*e*) Academic Achievement—setting of high standards; and (*f*) Vocational Climate—practical, applied activities. In other words, students found that Princeton had less of these qualities than they had anticipated. One area in which students tended to find more than the expected was in the press for Student Dignity—student freedom and allowance for personal responsibility. In sum, students generally expected Princeton to be characterized by more of an intellectual climate and more of a nonintellectual climate than they found it to be.

Where do these expectations come from? Some may be characteristic of all students entering college. However, in an attempt to determine whether certain characteristics were associated with one or another set of expectations, means on the scales and factors were computed for the following groups of students: public high school—private high school, alumni sons—nonalumni sons, below Scholastic Aptitude Test-Verbal (SAT-V) mean—above SAT-V mean, below Activities Index Area 1 (intellectual orientation) mean—above Activities Index Area 1 Mean, and AB students—BSE (engineering) students. This analysis resulted in many significant differences on the scales—22 for type of high school, 19 for type of son, 16 for Activities Index mean, eight for SAT-V mean, and four for type of degree candidate. For example, students from private high schools expected Princeton to have more press for both intellectual climate and nonintellectual climate than did students from public schools, and students with SAT-V scores below the mean expected less press for intellectual climate but more press for nonintellectual climate than students with scores above the SAT-V mean. These data suggest that demographic, personality, and intellectual variables all play a part in determining student expectations of Princeton, and very likely of other colleges as well.

In regard to student self-expectations, the data reported here indicate that many talented freshmen at Princeton overestimate their future level of academic performance and set high standards that they feel they must meet. Similarly, Berger (1) found that at the University of Chicago 67 per cent of the entering students expected to make Dean's List, whereas only 19 per cent would in fact do so. Student self-expectations at Princeton do not seem merely to reflect predictions based on past performance, since after one year at Princeton 85 per cent of the students report that they expect to stand in the top half of the class upon graduation, and 80 per cent give the top half as the minimum acceptable level.

While students thus appear to be fairly inaccurate in their estimates of their future academic performance, it is worthwhile noting that expected decile upon graduation correlated .29 ($p < .001$) with first semester average and with average for the first year. The corresponding correlations for SAT-V are .36 and .34, respectively. Thus the self-expectations appear to be quite inaccurate in absolute terms, but somewhat accurate when the relative performance of the students is considered. This may correspond to the finding that expectations were quite accurate for the relative strength of the presses at Princeton, but quite inaccurate for the absolute strength.

The data appear to be accurate assessments of the expectations of students entering Princeton. Use of the University Health Services as the sponsor of the research, assurances of confidentiality, and the voluntary nature of the project appear to have helped the students to feel free to respond in an honest manner. The return of over 80 per cent of the CCI questionnaires indicates that a reasonable proportion of the class was sampled.

There remains some question as to how these expectations might best be characterized. Should they be viewed as unrealistic, nonrealistic, inaccurate, or as fantasies? The data have indicated that the expectations have both accurate and inaccurate components. Furthermore, it may be that it is not unrealistic for over 80 per cent of the class to expect to stand in the top half of the class, since past performance justifies high aspirations in most of these students. On the other hand, the minimum acceptable levels of performance are also quite high, the self-expectations do not change much after one year, and the expectations concerning Princeton appear to have an optimistic or fantasy quality to them.

It may be that some expectations are based on past performance or inaccurate information, and these are quickly subject to change once the actual college situation is encountered. Other expectations may be a function of the needs of the individual and be quite resistant to change. This latter possibility raises the following question: What are the effects upon the student of disconfirmations of his expectations? It is likely that the effects will be influenced by the extent of the disconfirmation, the nature of the disconfirmation, and the way in which the individual copes with surprise, uncertainty, and disappointment in his environment. Current plans are to relate discrepancies between expected and actual press to the academic performance, health, and satisfaction with college of individual students.

E. SUMMARY

Students entering Princeton University indicated their expectations about the Princeton environment by responding to the College Characteristics Index prior to arrival on campus. Expectations about their own performance at college were indicated during freshman orientation week. Expectations about the press of the Princeton environment were accurate in terms of the relative strength of the presses, but inaccurate in terms of the absolute strength of the presses. Similarly, there was a tendency for self-expectations to be relatively associated with actual academic performance, but students generally overestimated their level of performance. Demographic, intellectual, and personality variables were all related to differences in expectations.

REFERENCES

1. BERGER, A. The unexpected findings: A study of the expectations of college students. Unpublished manuscript, National Opinion Research Center, University of Chicago, Chicago, Illinois, 1963.

2. CARLSMITH, J. M., & ARONSON, E. Some hedonic consequences of confirmation and disconfirmation of expectancies. *J. Abn. & Soc. Psychol.*, 1963, **66**, 151-156.
3. PERVIN, L. A. The need to predict and control under conditions of threat. *J. Personal.*, 1963, **31**, 570-587.
4. STERN, G. G. Environments for learning. In N. Sanford (Ed.), *The American College*. New York: Wiley, 1962. Pp. 690-730.
5. THISTLETHWAITE, D. L. College press and change in study plans of talented students. *J. Educ. Psychol.*, 1960, **51**, 222-234.
6. WEBB, S. C. Eight questions about the Emory environment. Testing and Counseling Service Memorandum, Emory University, Emory University, Georgia, September, 1963.

Department of Psychology
Green Hall
Princeton University
Princeton, New Jersey 08540

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 49-57.

THE INTERPERSONAL VALUES OF PARENTS OF ACHIEVING AND NONACHIEVING GIFTED CHILDREN*

Department of Psychology, University of New Mexico

RALPH D. NORMAN

A. INTRODUCTION

Recent years have seen intensive investigation into the area of underachievement, with emphasis being given to scrutiny of society's most precious resource —its gifted children. The achievement motive itself has been the subject of much research by McClelland and his associates (9, 11, 12). In the rather elaborate framework McClelland provides for this motive, he points out that, in order for children to develop high achievement motivation, parents must provide the child with opportunities for mastery which, because they are just beyond his present knowledge or attainment, will provide continuing stimulation for him. McClelland also mentions that individuals may develop achievement frames of reference because of stress placed on them by parents. Other investigators (2, 6, 13, 14, 16, 19, 20, 22, 23) have pointed out how parental expectations, such as training for independence, and values held by parents, such as democratic or dominative attitudes, may be related respectively to achievement and underachievement. In his fourth volume on the gifted, Terman (21) contrasts a "most successful" group of adults with a "least successful" group, and careful reading of his data leaves the reader with some impression that home atmospheres in childhood had relationship to adult achievement. Drawing from a large number of studies of bright students, Gowan (7) has attempted to summarize the significant factors that contribute to achievement. He refers to the well known facts that socioeconomic status, educational level, community activity of the parents, etc., affect achievement, but also points out how consonance of parental and child values and parental involvement in demands made upon the child are influential.

Moreover, within the general framework of parent-child interaction, agreement is quite general that identification—a mechanism strongly stressed in psychoanalytic theory—is quite important. Generally, it has been noted that children tend to adopt the value orientations of the parent who is of the

---

* Received in the Editorial Office, Provincetown, Massachusetts, on June 21, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

same sex as themselves. Whether one prefers psychoanalytic or social role theory formulations, it would seem that boys generally introject the standards, ideals, and values of their fathers, and girls those of their mothers.

Thus it may be assumed that the values that parents, especially those of the same sex, hold, as well as the congruence of these values between parents, should have some effect on a child's achievement. The present study is an attempt to explore these values among parents of a group of gifted children differing as to achievement status.

B. Procedure

In connection with another study, the author administered the Gordon Survey of Interpersonal Values (5) to parents of gifted children who had been categorized respectively as achievers and nonachievers. Method of assignment of the children to either the achieving or nonachieving category is delineated at length in an earlier study by the author and his students (15). Briefly, these investigators started with all children in the sixth grade of the Albuquerque public schools who had an *IQ* of 130 or more on the California Test of Mental Maturity (CTMM). Using a method devised by Shanner (17) that contrasts a child's "anticipated" achievement on the California Achievement Test (CAT) derived from his CTMM score with his actual achievement, and using arbitrary cutting scores necessitated by sex differences in the actual-anticipated differential, the investigators were able to designate 27 achieving boys, 32 nonachieving boys, 22 achieving girls, and 18 nonachieving girls from an original group of 215 gifted children. There were thus 49 achievers and 50 nonachievers, with each group divided according to sex in roughly equivalent proportions.

The cooperation of most of the parents of these children was obtained in completing the Gordon test. There were 23 mother-father pairs who were parents of achieving boys, 25 pairs who were the parents of nonachieving boys, 19 pairs who were parents of achieving girls, and 15 pairs who were parents of nonachieving girls. These pairs were used in calculating the *r*'s reported below. (There were a few additional unpaired parents whose scores were also used in making similar-sex comparisons as later reported.) The percentages of total possible parents who took the Gordon test were as follows: mothers of achievers, 90 per cent; mothers of nonachievers, 90 per cent; fathers of achievers, 86 per cent; fathers of nonachievers, 82 per cent.

The Survey of Interpersonal Values (SIV) measures six dimensions: (*a*) Support: being treated with understanding, receiving encouragement from other people, being treated with kindness and consideration; (*b*) Conformity:

doing what is socially correct, following regulations closely, doing what is accepted and proper, being a conformist; (*c*) Recognition: being looked up to and admired, being considered important, attracting favorable notice, achieving recognition; (*d*) Independence: having the right to do whatever one wants to do, being free to make one's own decisions, being able to do things in one's own way; (*e*) Benevolence: doing things for other people, sharing with others, helping the unfortunate, being generous; and (*f*) Leadership: being in charge of other people, having authority over others, being in a position of leadership or power (5, p. 3).

The questions the investigator intended to answer with the SIV center around the problems of identification of the child with the same-sexed parent and that of congruence between parental values. First, in order to test the effects of the values of the parents of the same or opposite sex on the achievement orientations of their children, the following two hypotheses were formulated: (*a*) there will be no significant differences between comparable mean value scores on the SIV of parents of either category of gifted children who are of the same sex as the child (this means contrasting mean scores of fathers of achieving boys with those of nonachieving boys, and those of mothers of achieving girls with those of mothers of nonachieving girls); and (*b*) there will be no significant differences between comparable mean value scores of parents who are of the opposite sex of the child who falls into either category (this means contrasting mean scores of fathers of achieving girls with those of fathers of nonachieving girls, and those of mothers of achieving boys with those of mothers of nonachieving boys). Second, in order to assess the comparability of relationship between mother-father pairs who are parents of either category of achieving child, the following three hypotheses were formulated: (*c*) when correlations are obtained between mother-father pairs for the separate value scores, there will be no variation in the *pattern* of significant correlations obtained; (*d*) there will be no significant differences in comparable correlations when these are tested between "achieving" and "nonachieving" parental pairs; and (*e*) there will be no significant differences in the variability of these correlations between the two groups.

C. RESULTS

Results of the study are presented in Tables 1 and 2. Table 1 gives Gordon SIV mean scores of like-sexed parents of either category of gifted child, presenting data on a direct test of the process of same-sex identification. Of the six values of the SIV, the mean scores for two, Independence and Conformity, show significant or close to significant differences. (The *t* of 1.94 for

TABLE 1

GORDON SIV SCORES OF LIKE-SEXED PARENTS OF ACHIEVING AND NONACHIEVING GIFTED CHILDREN

Group	S		C		R		I		B		L	
	M	σ	*M*	σ	*M*	σ	*M*	σ	*M*	σ	*M*	σ
						Fathers of boys						
Achievers ($N = 23$)[a]	15.48	4.14	13.74	5.18	10.91	4.77	17.74	6.23	15.61	5.60	16.35	8.87
Nonachievers ($N = 26$)	16.46	5.26	17.46	6.44	11.69	4.93	11.23	6.68	18.31	5.85	14.73	7.23
Difference	0.98		3.72		0.78		6.51		2.70		1.62	
t	0.70		2.16**		0.55		3.44***		1.61		0.69	
						Mothers of girls						
Achievers ($N = 20$)	19.45	5.05	15.70	5.51	9.45	3.57	18.85	5.70	20.70	3.55	5.85	3.61
Nonachievers ($N = 16$)	17.63	2.64	19.80	5.69	9.69	2.59	14.69	6.84	21.69	3.44	6.50	3.78
Difference	1.82		4.10		0.24		4.16		0.99		0.65	
t	1.27		2.04**		0.22		1.94*		0.82		0.51	

Note: S = Support, C = Conformity, R = Recognition, I = Independence, B = Benevolence, and L = Leadership.

[a] *N*s are those of parents, not of children.

* Significant between .05 and .10 levels.

** Significant at .05 level.

*** Significant at .01 level.

Independence in the case of the mothers falls just short of significance at the .05 level; *t* must be 2.03 for 34 *df*.) The fathers of boy achievers and the mothers of girl achievers have *higher* Independence scores but *lower* Conformity scores than do the comparable parents of the nonachievers. It is indeed rather striking that the identical values appear important in the case of ***both*** groups of same-sexed parents, and the investigator must reject the first hypothesis at least in the case of these two variables.

TABLE 2
CORRELATIONS (AND CORRESPONDING z'S) BETWEEN SIV SCORES OF MOTHERS AND FATHERS OF BOTH GROUPS

Group	S	C	R	I	B	L
				r's		
Achievers ($N = 42$)[a]	.341*	.275	.362*	.256	.342*	.238
Nonachievers ($N = 40$)	.044	.360*	.075	.416**	.444**	—.228
				z's		
Achievers ($N = 42$)	.355	.283	.379	.262	.356	.243
Nonachievers ($N = 40$)	.044	.377	.075	.443	.477	—.232
$z_1 - z_2$	.311	.094	.304	.181	.121	.475
$\frac{z_1 - z_2}{\sigma(z_1 - z_2)}$	1.36	0.41	1.32	0.79	0.53	2.07*

Note: S = Support, C = Conformity, R = Recognition, I = Independence, B = Benevolence, and L = Leadership.

[a] *N*s are those of husband-wife pairs, not of children.

* Significant at .05 level.

** Significant at .01 level.

When opposite-sex identification was examined by means of testing the second hypothesis, there were but two significant or almost significant differences, that in Leadership and Recognition only in the case of the mothers of the two categories of boys. (The *t* for one of these, Leadership, was significant between the .05 and .10 levels.) The mothers of achieving boys, compared with those of nonachieving boys, showed higher mean scores for Leadership, but lower mean scores for Recognition. However, absolutely no significant differences appeared in the case of contrasts of fathers of achieving girls with those of nonachieving girls. Rejection of the second null hypothesis may then be made with a good degree of confidence for only one variable, Recognition, in the case of the mothers of boys. It is difficult to conclude from the data that cross-sex identification is generally taking place.

Table 2 presents the correlations and corresponding z transformations, used to test the significance of the differences between the *r*'s for our fourth hypothesis. It is obvious that there is a pattern of significant correlations between parent scores of achievers different from the pattern between parent scores of nonachievers. The former have significant relationships for Support, Recognition, and Benevolence; the latter demonstrate significant relationships for Conformity, Independence, and Benevolence. The investigator must accordingly reject the third hypothesis, which states that there will be no differences in the *pattern* of significant correlations obtained. When the r's are transformed into z's and the latter tested for significance between the two groups, it may be seen that there is a significant difference between the two correlations for the Leadership variable only. The fourth hypothesis may be rejected then only in the case of this one value.

The investigator was led to test the fifth hypothesis by examining the z transformations. The latter are notable because of the comparatively small variability of the z's of the "achieving" parents across the board, and their clustering between a narrow range of about .24 to .38. The z's of the "nonachieving" parents, however, range between about —.23 and .48. The *SD* of the z's for the "achieving" parents was .052; that for the "nonachieving" parents was about five times as large, or .256. The $D/\sigma d$ was 2.48, significant between the .02 and .05 levels. Thus, despite the fact that only one z proved to be significantly different *between* the two groups, the "*within* group" variabilities of the z's are highly significantly different. The overall evidence from testing of the hypotheses concerning the correlations between husband-wife pairs is that there are important differences both in the patterns of correlations and in their variability between the two kinds of parents. Actually, the investigator can reject the last null hypothesis with a high degree of confidence.

D. Discussion

Since the SIV, according to Gordon (5) is "designed to measure certain critical values involving the individual's relationships to other people and their relationships to him," and since there is wide acceptance of the concept of identification, it is probably not surprising that like-sexed parents of achieving children show significantly higher scores in Independence and lower scores in Conformity than do like-sexed parents of nonachievers. It would appear that these 11-year-old children have already introjected certain value orientations from like-sexed parents. In terms of his relationships to other people, the "achieving" parent (with whom the child identifies) as contrasted with the "nonachieving" parent is one who values making his own decisions above that of conforming.

It is of special interest that significant differences are found for the values of Independence and its negatively related opposite, Conformity. (Gordon reports a correlation of —.38 between these two values among college students.) It would seem, according to previous research, that there is a linkage between what might be termed "independence" demands or values of parents and achievement motivation of children. Winterbottom (23) found that the earlier a child was expected to achieve independence, the higher was his need for achievement (nAch) score. In a cross-cultural study, McClelland and Friedman (10) found a relationship between earliness and severity of training for independence and nAch, although this finding was disputed by Child, Storm, and Veroff (4). Studying 9- to 11-year-old boys, one-half with high nAch and one-half with low nAch, Rosen and d'Andrade (16) had their subjects work on tasks under varying conditions designed to provide measures of parental socialization practices with respect to achievement training, independence training (including autonomy and self-reliance), and sanctions. Parents of boys high in nAch set high standards for them, expecting them to do well. Rosen and d'Andrade feel, however, that a distinction must be made between a general encouragement of independence in children on the one hand and the more specific fostering of independent achievement on the other. More recently, Shaw (18) contrasted the attitudes of parents of bright academic achievers with those of underachievers, using Winterbottom's Independence Training Inventory (23). Among other findings, he reports that parents of achievers wanted their children to learn to make their own decisions and expected their children to be more adult in their behavior at an earlier age than did parents of nonachievers. However, findings opposite to Shaw's are reported by Chance (3), who found that bright first graders whose mothers favored early independence training made less adequate school progress relative to their *IQ* than did children whose mothers favored later independence training. Chance does comment though that a mother's *attitudes* toward early independence training will differentially influence the child, depending as well upon the degree of her closeness and warmth to him.

Regarding the role of identification particularly in achievement motivation, Kagan and Moss (8) report a positive relationship between early maternal concern with achievement and both achievement mastery and *IQ* gain in girls, although they found no such relationship for boys. They also report a suggestive relationship between achievement fantasy in children and the educational level of the same-sexed parent. Argyle and Robinson (1) found that achievement motivation in grammar school children was found to be correlated with the reported achievement tendencies of the same-sexed parent and with identification with parents, provided that the identification was sufficient. Thus, it would seem that if nAch in children is associated with achievement strivings in the parent (especially the same-sexed one) and the latter in turn is a reflection of, or is somehow associated with, an independence need, it is reasonable to find the parents of the achieving gifted children higher in Independence and lower in Conformity on the Gordon SIV than parents of nonachievers.

Regarding the second group of findings—that concerning the *r*'s between mother-father pairs—it is difficult to find explanations because of paucity of previous research. The results, showing that there is a different pattern of significant correlations for the six value scores of the SIV, seem to indicate a different organization of value congruences between the parental pairs of both groups of children. However, in the opinion of the investigator, this finding is not as important as the fact that there is more consistency of relationship in values (significantly less variability in the *r*'s) between father-mother pairs of achievers than of nonachievers. Does this greater uniformity of value relationships denote increased agreement or harmony between achievers' parents? What does this signify for the performance of their children? It is of some interest here that in earlier research with the same children (15) the investigators found that the children themselves showed important differences in consistency of their test performance. Achievers were much more consistent than nonachievers both in their mean scores on the two parts of the CTMM (language and nonlanguage) and in their anticipated achievement profiles on the CAT. Whether greater consistency in parental values for achievers' as compared with nonachievers' parents is somehow productive of this sort of performance in their children would have to be ascertained by future research.

E. Summary

Parents of gifted children (*IQ* 130 or more) who had been categorized respectively as achievers or nonachievers were given the Gordon Survey of Interpersonal Values (SIV). It was found that the identification process was strongly operative: fathers of achieving boys and mothers of achieving girls made significantly higher mean scores in Independence and lower mean scores in Conformity than did the same-sexed parents of nonachievers. When correlations were run between father-mother pairs, it was found that different patterns of significant correlations on the SIV emerged for the respective parent groups, and the correlations for parents of achievers were significantly less variable than those for parents of nonachievers.

References

1. Argyle, M., & Robinson, P. Two origins of achievement motivation. *Brit. J. Soc. & Clin. Psychol.*, 1962, **1**, 107-120.
2. Barrett, H. O. An intensive study of 32 gifted children. *Person. & Guid. J.*, 1958, **36**, 192-194.
3. Chance, J. E. Independence training and first graders' achievement. *J. Consult. Psychol.*, 1961, **25**, 149-154.

4. CHILD, I. L., STORM, T., & VEROFF, J. Achievement themes in folk tales related to socialization practice. In J. W. Atkinson (Ed.), *Motives in Fantasy, Action and Society*. Princeton, N. J.: Van Nostrand, 1958. Pp. 479-492.

5. GORDON, L. V. SRA Manual for Survey of Interpersonal Values. Chicago: Sci. Res. Assoc., 1960.

6. GOWAN, J. C. Dynamics of the underachievement of gifted students. *Exceptional Child.*, 1957, **24**, 98-122.

7. ———. Factors of achievement in high school and college. *J. Counsel. Psychol.*, 1960, **7**, 91-95.

8. KAGAN, J., & MOSS, H. A. Stability and validity of achievement fantasy. *J. Abn. & Soc. Psychol.*, 1959, **58**, 357-364.

9. MCCLELLAND, D. C. (Ed.). Studies in Motivation. New York: Appleton-Century-Crofts, 1955.

10. MCCLELLAND, D. C., & FRIEDMAN, G. A. A cross-cultural survey of the relationship between child-training practices and achievement motivation appearing in folk tales. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in Social Psychology*. New York: Holt, 1952. Pp. 243-249.

11. MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.

12. MCCLELLAND, D. C., BALDWIN, A. L., BRONFENBRENNER, U., & STRODTBECK, F. L. Talent and Society. Princeton, N. J.: Van Nostrand, 1958.

13. MCDONALD, W. T. The relationship of achievement motivation to independence training, parental attitudes, ecology, and other factors. Unpublished Doctoral dissertation, University of Houston, Houston, Texas, 1956.

14. MORROW, W. R., & WILSON, R. C. Family relations of bright high-achieving and under-achieving high school boys. *Child Devel.*, 1961, **32**, 501-510.

15. NORMAN, R. D., CLARK, B. P., & BESSEMER, D. W. Age, sex, IQ, and achievement patterns in achieving and nonachieving gifted children. *Exceptional Child.*, 1962, **29**, 116-123.

16. ROSEN, B. C., & D'ANDRADE, R. The psychological origins of achievement motivation. *Sociometry*, 1959, **22**, 185-218.

17. SHANNER, W. M. New concepts in norms. In A. E. Traxler (Ed.), *Positive Values in the American Educational System*. Washington, D.C.: Amer. Council on Educ., 1958. Pp. 64-74.

18. SHAW, M. C. Note on parent attitudes toward independence training and the academic achievement of their children. *J. Educ. Psychol.*, 1964, **55**, 371-374.

19. SHAW, M. C., & DUTTON, B. E. The use of the Parent Attitude Research Inventory with the parents of bright academic under-achievers. *J. Educ. Psychol.*, 1962, **53**, 203-208.

20. TEAHAN, J. E. Parental attitudes and college success. *J. Educ. Psychol.*, 1963, **54**, 104-109.

21. TERMAN, L. M., & ODEN, M. H. Genetic Studies of Genius: Volume IV. The Gifted Child Grows Up. Stanford, Calif.: Stanford Univ. Press, 1947.

22. WEIGAND, G. Adaptiveness and the role of the parents in academic success. *Person. & Guid. J.*, 1957, **35**, 518-522.

23. WINTERBOTTOM, M. R. The relation of childhood training in independence to achievement motivation. Unpublished Doctoral dissertation, University of Michigan, Ann Arbor, Michigan, 1953.

Department of Psychology
University of New Mexico
Albuquerque, New Mexico

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 59-61.

THE RETENTION OF SENSORY EXPERIENCE: VI. STIMULUS REPETITION AND INTERFERENCE EFFECTS*

University of Pittsburgh, School of Medicine

H. E. King

A. Problems

Human subjects have shown unexpected accuracy in retaining the stimulus properties of intensity, frequency, and duration for simplified percepts formed via visual or auditory modalities (1, 2, 3), over short-delay and long-delay intervals (4), and when a systematic variation of the sensory value has been explored (5). Two related questions concern the influence, if any, of (*a*) multiple exposure to the sensory stimulus to be "held" through time—as a repeated standard stimulus may permit a degree of correction in judged Point of Subjective Equality—and (*b*) interference by means of extraneous stimuli introduced during the period of delay to reproduction—as either inhibitive or associative interference might then occur.

B. Procedures

1. *Experiment 1*

Retained impressions of *Auditory Intensity* (*Loudness*) were chosen to sample possible differences in the accuracy of matching performance following single or multiple exposure to a *standard* stimulus, using a method previously described (1). Four matched groups of eight subjects each were exposed to a standard stimulus tone (600 cps; 43.8 db; 5 sec.s), sounded from one to four times, with a five-second delay between exposures. Following a two-minute delay, each subject attempted to match the standard by adjusting the intensity of a *variable* stimulus tone of the same frequency.

2. *Experiment 2*

The ability to reset *Auditory Frequency* (*Pitch*) stimuli to match a foregoing standard tone was examined by a method earlier described (2), following the interposition of extraneous auditory stimuli during the period of delay to reproduction. A *standard* stimulus tone (1000 cps; 75 db; 5 sec.s) was

---

* Received in the Editorial Office, Provincetown, Massachusetts, on June 21, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

matched by the adjustment of a *variable* stimulus tone after a two-minute delay, following conditions of (*a*) silence, (*b*) white noise, (*c*) higher tone (2000 cps, equated for intensity) and lower tone (500 cps, equated for intensity) interposed for 20 seconds beginning five seconds after the *standard* stimulus went off.

C. RESULTS

Accuracy in retaining was evaluated by the analysis of variance, comparing the delayed matches made under all experimental conditions. No statistically significant departures were observed in matching performance under any of the trial conditions in either experiment. Supplemental comparisons were also made, using Fisher's t test, of the grand mean of all matched adjustments of variable stimuli with the physical value of the relevant standard stimuli (invariant) and between the matching performance of individual conditions by subgroup. Repeated exposure of the subject to the (*Loudness*) standard stimulus, in Experiment 1, was singularly without influence on delayed matching performance ($F = 1.115$ where 5 per cent level is 8.62). All matches made were close to the intensity of the standard, falling equally and unsystematically above or below it ($t = 0.65$ where 5 per cent level is 2.042). A more noticeable effect was observed for delayed matching performance in Experiment 2, in which extraneous auditory stimuli were interposed after a standard (*Pitch*) stimulus ($F = 2.273$ where 5 per cent level is 8.62). A reduced group variability characterized all conditions of sound interposition; high and low tone were associated with nonsignificant reductions in judged equivalence, while white noise raised the averaged resetting to a level approaching statistical reliability. The difference in means between the subgroups (conditions) at the extreme of these opposed trends—high tone contrasted with white noise—did achieve significance ($t = 2.63$ where 5 per cent level is 2.365).

D. SUMMARY AND CONCLUSIONS

As none of the main conditions explored produced a reliable difference in matching performance, the series as a whole may be summarized readily: neither repeated exposure to a *standard* tone stimulus, nor the interposing of irrelevant sound during delay to reproduction, exerted noticeable influence on judgments of matched equivalence made after brief delay. Stated positively, these findings are in harmony with conclusions drawn in earlier experimentation with sensory retaining. That is to say, the accuracy with which simplified sensory percepts can be reproduced following brief or prolonged delay—*viz.,* *brightness, flash-rate, visual duration,* and the like (1, 2, 3)—has been shown to have its counterpart in the high degree of accuracy with which varied values of such standard stimuli can be matched following short delay (5). The findings of the two current experiments, although they represent a limited sampling only, would seem to indicate that the firmness of the "trace," or strength of the original impression, is such that it is expressed also in its resistance to influences that might be thought to either aid (repetition) or hinder it (interference). The opposing trends observed for conditions of high or low tone interference compared with white noise interposition may reflect the different anchoring influence of sound-tone as against sound-noise: *i.e.,* the interference effect of a sound with "meaning" (another pitch that relates to the identifiable pitch of the standard stimulus) as opposed to a sound stimulation with less identifiable "meaning" (white noise being a wide-frequency sound spectrum). This possible difference in anchoring calls for further exploration, as does the more general question of "trace" interference in sensory-retaining, for the present experiments serve more as a sample or guide to the interactions that may occur than as a broad scale demonstration of effect.

REFERENCES

1. KING, H. E. The retention of sensory experience: I. Intensity. *J. of Psychol.,* 1963, **56**, 283-290.
2. ———. The retention of sensory experience: II. Frequency. *J. of Psychol.,* 1963, **56**, 291-298.
3. ———. The retention of sensory experience: III. Duration. *J. of Psychol.,* 1963, **56**, 299-306.
4. ———. The retention of sensory experience: IV. Short-delay *versus* long-delay intervals. *J. of Psychol.,* 1965, **60**, 103-115.
5. ———. The retention of sensory experience: V. Variation of the standard stimuli. *J. of Psychol.,* 1966, **62**, 15-22.

Western Psychiatric Institute and Clinic
University of Pittsburgh
School of Medicine
3811 O'Hara Street
Pittsburgh, Pennsylvania 15213

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 63-68.

SCHOLASTIC APTITUDE AND SMOKING ATTITUDE AND BEHAVIOR OF COLLEGE MALES*

Department of Psychology, Boston College

DANIEL J. BAER

A. INTRODUCTION

So far many factors have been associated with an individual's smoking attitude, behavior, and beliefs (10). Although smokers differ from nonsmokers in school achievement (8), other variables, such as education level (7) and *IQ* (9), have not distinguished these two groups.

The purpose of this study was to compare college male smokers and nonsmokers in their Scholastic Aptitude Test Verbal (SAT-V) and Mathematics (SAT-M) performance, and to examine the relationship between SAT scores and the smoking attitude of these groups. Since habitual smoking is physically harmful (13) and thus undesirable, it would seem that brighter students would more likely avoid smoking and would have a more unfavorable attitude toward smoking. As a working hypothesis it was suggested that smokers would tend to have lower SAT-V and SAT-M scores than nonsmokers, and that the higher an individual's SAT scores the more unfavorable would be his attitude toward smoking.

B. METHOD

1. *Subjects*

Only white, Catholic, American-born, male college students were selected as *S*s. The 249 Present Smokers were habitual smokers at the time they were tested, the 32 Former Smokers were habitual smokers at one time but had not smoked for at least one month prior to the testing session, and the 124 Nonsmokers had never smoked except for a possible rare trial experience. The three groups were comparable in age and number of years and selected major in college.

2. *Test Material and Procedure*

The Smoking Attitude Scale (SAS), which has been described in a previous study (2), was a Thurstone-type attitude scale consisting of two sets of 22

---

* Received in the Editorial Office, Provincetown, Massachusetts, on June 27, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

statements. SAS scores were based upon the sum of the weighted values of these statements to which the *S*s expressed agreement. Higher scores represented more favorable smoking attitudes. Twenty male college students secured the 405 *S*s from many sources on campus.

The SAT-V and SAT-M scores, which represent general academic ability, were obtained from the *S*s' admission records.

C. Results and Discussion

Table 1 summarizes the SAT-V and SAT-M means and standard deviations for the Present Smoker, Former Smoker, and Nonsmoker groups, and the *t*-tests of the difference between the means for the groups. It may be seen that, for the one-tailed hypothesis, Former Smokers had significantly (.05 level) lower SAT-V scores than did the Present Smoker and Nonsmoker groups, but no difference was found between the Present Smoker and Nonsmoker groups. Although such a finding was contrary to the predicted outcome, the acceptance of this result may not merely represent a type I error. Students of lower scholastic aptitude may be more prone to change than students of higher aptitude. Since in a previous study (2) the smoking attitude of Former Smokers was more unfavorable than that of the Present Smokers, and since 100 per cent of them reported a belief in the relationship between lung cancer and smoking, it may be that these students had more readily accepted the reports concerning the harmful effects of smoking. It is interesting that the mean SAT-V score for the Present Smokers and Nonsmokers corresponded to the mean level of the college population from which they were sampled (the mean for SAT-V was 545 and for SAT-M was 555); however, the mean performance of the Former Smoker group was atypically lower than that of the other groups. It seems as though lower

TABLE 1
t Test of Difference Between Mean Scholastic Aptitude Test Performance of Present Smoker, Former Smoker, and Nonsmoker Groups

Scholastic Aptitude Test	Group	*N*	Mean	*SD*	Former smoker	Non-smoker
					t	
Verbal	Present smokers	249	543.6	91.1	1.81*	0.39
	Former smokers	32	513.2	75.8		—1.92*
	Nonsmokers	124	547.6	93.5		
Mathematics	Present smokers	249	550.8	86.4	0.52	1.74*
	Former smokers	32	559.4	91.0		0.46
	Nonsmokers	124	567.3	85.3		

* $p < .05$ (one-tailed hypothesis).

aptitude Present Smokers are more likely candidates for giving up the smoking habit.

The significantly (.05 level) lower Present Smokers' SAT-M scores, when compared to the Nonsmokers' scores, were consistent with the predicted outcome. Although the 17-point spread may be a small practical difference, the mean SAT-M score of the Nonsmokers was not only higher than that of the Present Smokers, but was higher than that of the population mean from which this sample was selected. This difference may be partially explained on the basis of the type of major selected by the students. Whereas 17 per cent of the Present Smokers majored in the natural sciences, 25 per cent of the Nonsmokers had done so. Physical science majors would more likely do better on the SAT-M test, and also are less likely to smoke than other students (5).

The significant difference between the means in Table 2 indicates that, for the Present Smokers, the heavy smokers had higher SAT-V scores than did light or moderate smokers, and that the students who smoked only cigarettes had lower SAT-V and SAT-M scores. When compared to Table 1 results, as well as the college population means, the "not heavy smoker" group more closely resembled the Former Smoker than the Nonsmoker group. Although people offer different reasons for why they begin smoking (11, 12), so far the factors contributing to the onset of heavy smoking have not been extensively explored. The present outcome at least demonstrates that, besides their difference in personality traits (5, 6, 11), heavy smokers differ from light and moderate smokers in verbal aptitude. The opposite trend apparent for the SAT-M scores may reflect the greater susceptibility of these scores

TABLE 2

t TEST OF DIFFERENCE BETWEEN SCHOLASTIC APTITUDE TEST PERFORMANCE FOR PRESENT SMOKERS PRACTICING DIFFERENT SMOKING BEHAVIOR

Scholastic Aptitude Test	Present smokers	*N*	Mean	*SD*	*t*
Verbal	Heavy smokers	143	552.4	89.0	1.76*
	Not heavy smokers	106	531.9	93.5	
	Smoke cigarettes only	138	529.1	85.5	2.85**
	Smoke pipe or cigars	111	561.7	96.0	
Mathematics	Heavy smokers	143	546.4	82.9	0.94
	Not heavy smokers	106	556.3	91.0	
	Smoke cigarettes only	138	539.9	77.2	2.23*
	Smoke pipe or cigars	111	564.3	95.0	

* $p < .05$ (one-tailed hypothesis).

** $p < .01$ (one-tailed hypothesis).

to educational experiences. Perhaps a common factor, such as the constitutional hypothesis proposed by Eysenck and others (4), could account for the heavy smoking and high verbal ability of some students.

Many personal and social conditions influence some men to smoke only cigarettes, and others to smoke only a pipe or cigars. Eysenck (3) reports that pipe smokers are the most introverted of all other groups of smokers and nonsmokers. Perhaps a role-taking behavior based upon stereotyped images associated with pipe or cigar smoking may have motivated some students to assume this form of smoking activity. Since 61 per cent of the social science majors smoked cigarettes only, while 65 per cent of the natural science majors smoked a pipe or cigars, it would seem that major area of study in college represents an important correlate of type of smoking behavior.

The negative Pearson correlation coefficients reported in Table 3 indicate that the higher the SAS scores, the lower the SAT-V and SAT-M scores.

TABLE 3
CORRELATION BETWEEN SCHOLASTIC APTITUDE TEST AND SMOKING ATTITUDE SCALE SCORES FOR PRESENT SMOKERS, FORMER SMOKERS, AND NONSMOKERS

	Group		
Scholastic Aptitude Test	Present smokers ($N = 249$)	Former smokers ($N = 32$)	Nonsmokers ($N = 124$)
Verbal	—.05	—.22	—.12
Mathematics	—.01	—.31*	—.21*

* $p < .05$ (one-tailed hypothesis).

Although only two of the correlations were significant (.05 level, one-tailed test), all of the values were in a direction indicating that brighter students tended to have more favorable smoking attitude. This result is surprising in view of the significantly less favorable smoking attitude held by the Former Smoker and Nonsmoker groups when compared to the Present Smoker groups (2). Perhaps a more rational, intellectual acceptance of smoking can be entertained by brighter students who are not actually victims of the habit.

Since attitudes toward mathematics influence performance on mathematics aptitude tests (1, 9), some common environmental influence may have indirectly contributed to the favorable smoking attitude and high mathematical aptitude of these Former Smokers and Nonsmokers. It may be that these students are prone to give more extreme responses to attitude scales. It seems unlikely that the SAS is a more direct measure of scholastic aptitude, especially since significant correlation was not found for the SAT-M scores. A biographical search for the causes of these correlations would be an interesting area for future research.

In Table 4 are reported the Pearson correlation coefficients between the SAS and SAT scores for Present Smokers who differ in their smoking behavior. The significant (.05 level) correlations indicate that brighter students who smoke a pipe or cigars also have a more favorable smoking attitude. Since a similar result was not found for those students who smoke only cigarettes, it would seem that the motivating factors that attract brighter students to smoke a pipe or cigars are different from the factors that attract the brighter students to smoke cigarettes only. It is interesting that, although insignificant, the brighter heavy smokers and those who smoke only cigarettes tended to have a more unfavorable smoking attitude.

TABLE 4
CORRELATION BETWEEN SCHOLASTIC APTITUDE TEST AND SMOKING ATTITUDE SCALE SCORES FOR FOUR GROUPS OF PRESENT SMOKERS

Scholastic Aptitude Test	Heavy smokers ($N = 143$)	Not heavy smokers ($N = 106$)	Smoke cigarettes only ($N = 138$)	Smoke pipe or cigars ($N = 111$)
Verbal	.09	—.01	.08	—.14*
Mathematics	—.03	—.06	—.13	—.16*

* $p < .05$ (one-tailed hypothesis).

D. SUMMARY

A sample of 249 Present Smokers, 32 Former Smokers, and 124 Nonsmokers (all male college students) were compared in their Scholastic Aptitude Test performance and in the relationship between these scores and their smoking attitude. Nonsmokers and Present Smokers who smoked a pipe or cigars tended to have higher mathematical aptitude than did other students, while Former Smokers and light or moderate smokers tended to have lower verbal aptitude. In general, a positive relationship was observed between high scholastic aptitude and favorable smoking attitude.

REFERENCES

1. AIKEN, L. R., JR., & DREGER, R. M. The effect of attitudes on performance in mathematics. *J. Educ. Psychol.*, 1961, **52**, 19-24.
2. BAER, D. J. Smoking attitudes, behavior, and beliefs of college males. *J. Soc. Psychol.*, 1966, **68**, 65-78.
3. EYSENCK, H. J. Personality and cigarette smoking. *Life Sci.*, 1964, **3**, 777-792.
4. ———. Smoking, Health, and Personality. New York: Basic Books, 1965.

5. HEATH, C. W. Differences between smokers and nonsmokers. *AMA Arch. Intern. Med.,* 1958, **101**, 377-388.
6. LAWTON, M. P., & PHILLIPS, R. W. The relationship between excessive cigarette smoking and psychological tension. *Amer. J. Med. Sci.,* 1956, **232**, 397-402.
7. LILIENFELD, A. M. Emotional and other selected characteristics of cigarette smokers and nonsmokers as related to epidemiological studies of lung cancer and other diseases. *J. Nat. Cancer Inst.,* 1959, **22**, 259-282.
8. LYNN, R. M. A study of smokers and nonsmokers as related to achievement and various personal characteristics. University of North Carolina Records of Research Progress, No. 464 (Graduate School Series No. 56), 1948 (Abstract).
9. MATARAZZO, R. G., MATARAZZO, J. D., & PHILLIPS, J. S. Psychological test and organismic correlates of interview interaction patterns. *J. Abn. & Soc. Psychol.,* 1958, **56**, 329-338.
10. MATARAZZO, J. D., & SASLOW, G. Psychological and related characteristics of smokers and nonsmokers. *Psychol. Bull.,* 1960, **57**, 493-513.
11. SALBER, E. J., & ROCHMAN, J. E. Personality differences between smokers and nonsmokers. *Arch. Environ. Health,* 1964, **8**, 459-465.
12. SALBER, E. J., WELSH, B., & TAYLOR, S. V. Reasons for smoking given by secondary school children. *J. Health & Hum. Behav.,* 1963, **4**, 118-129.
13. THE SURGEON GENERAL'S ADVISORY COMMITTEE ON SMOKING AND HEALTH. Smoking and Health. Washington, D.C.: U.S. Gov. Printing Office, 1964.

Department of Psychology
Boston College
Chestnut Hill, Massachusetts 02167

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 69-71.

A SYNTHESIS OF PERSONALITY THEORIES BY CLUSTER ANALYSIS*

Berkeley, California

ALLEN J. SCHUH

A. INTRODUCTION

Comparisons of personality theories are frequently beyond the comprehension of students in psychology. Although a synthesis for review purposes was provided in Hall and Lindzey's *Theories of Personality* (1, p. 548), it is felt that these authors could have gone several steps further (*a*) by reducing the variables on which they rated the theorists to basic dimensions by cluster analysis, and (*b*) by reducing the theories themselves to basic clusters. This brief note reports the results of performing these analyses.

B. PROCEDURE

In rating the theorists, Hall and Lindzey used three categories: high, medium, and low. The ratings depended upon whether the theory emphasized, occupied a middle ground, or de-emphasized the importance of the issue or set of determinants. High ratings were converted to a value of one, moderates to two, and lows to a value of three. A Cumulative Communality Key Cluster Analysis (2) was performed for both the variable (V) and theorist (O) analyses. The theorists were assigned simple sum T score cluster scores on the orthogonal dimensions resulting from the V-analysis. The correlation matrix was then transposed and an O-analysis performed. The resulting orthogonal dimensions were rotated to oblique structure. Three clinical psychologists assisted in naming the resulting clusters.

C. RESULTS

1. The first of four dimensions resulting from the V-analysis was composed of the items: purpose, psychological environment, self concept, and field emphasis; represented by the factor coefficients .96, .83, .67, and .66, respectively. This dimension refers to the *usefulness* of the self concept in changing behavior. The second dimension was composed of the items: early develop-

---

* Received in the Editorial Office, Provincetown, Massachusetts, on June 29, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

mental experience, social science, reward, and continuity of development; represented by the factor coefficients .97, .84, .62, and .53, respectively. This dimension refers to the relative *importance* of the self concept in changing behavior. The third dimension was composed of the items: heredity, biology, personality structure, and social science; represented by the factor coefficients .83, .76, .70, and —.31, respectively. This dimension refers to the amount of change in behavior that is seen as possible by the theorists. The fourth and last dimension was composed of the items: learning process, contiguity, multiplicity of motives, and unconscious determinants; represented by the factor coefficients .74, .66, .55, and —.49, respectively. This dimension refers to the complexity of the process of change as seen by the theorists.

2. The cluster scores for the theorists are presented in Table 1 in simple sum T score form. Table 1 is presented as a teaching and review aid. The theorists may either be compared directly from the table or their T scores may be plotted. For example, it can be seen that Rogers, Goldstein, and Angyal appear at the one extreme (in contrast to Miller and Dollard, Eysenck, and Sheldon) on the issue of the usefulness of the self concept in changing behavior.

TABLE 1
PERSONALITY THEORIST'S CLUSTER SCORES ON FOUR DIMENSIONS

	Dimension			
Theorist	Usefulness of self concept	Importance of self concept	Amount of change possible	Complexity of the process
Freud	50	38	41	60
Jung	53	57	37	61
Adler	42	43	53	60
Horney	46	46	65	60
Fromm	50	46	61	55
Sullivan	38	43	57	40
Lewin	42	58	62	41
Allport	49	65	42	36
Murray	42	39	49	57
Sheldon	69	57	37	60
Eysenck	69	57	41	55
Cattell	57	49	37	36
Miller and Dollard	69	35	65	40
Angyal	42	57	46	55
Goldstein	42	65	50	49
Rogers	42	61	63	55
Murphy	46	35	44	31

Note: Scaling Constants: Mean = 50, $SD = 10$.

3. The first of four clusters of theorists evolving from the O-analysis consisted of Fromm, Horney, and Adler. Their factor coefficients were .92,

.79, and .66, respectively. This cluster is named Social Emphasis. The second cluster was composed of Goldstein, Angyal, Miller and Dollard, Allport, Rogers, and Murphy with factor coefficients of .78, .75, —.74, .73, .67, and —.30, respectively. This cluster is named Self Emphasis. Goldstein, Angyal, Allport, and Rogers have theories far more in agreement with one another than heretofore anticipated. These theories appear to be interchangeable when basic philosophies and not semantics are considered. The third cluster was composed of Eysenck, Sheldon, Freud, Lewin, and Murray with factor coefficients of .79, .75, .56, —.47, and .40, respectively. This cluster is less interpretable than the first and second, and no attempt is made to identify it. The fourth and last cluster of theories was composed of Cattell, Jung, and Sullivan with factor coefficients of .75, 72, and —.68, respectively. This cluster does not lend itself easily to interpretation and is not named.

Presented in the form of cluster scores, some of the theories are seen to be in combinations with other theories, which most students may not have anticipated because of the confounding of semantics with basic issues. Finding Sullivan's theory in a cluster with Cattell and Jung, but not in the Social Emphasis cluster, for example, is contrary to expectations. From Table 1 it can be seen that the greatest discrepancy between Sullivan and the Social Emphasis theories is on the issue of the complexity of the process of change. Finding Cattell (a so-called factor theorist) in a cluster with Jung and Sullivan, but not with Eysenck (another factor theorist), indicates that their theoretical frameworks are further removed than their methods of investigation.

D. CONCLUSIONS

Despite the many apparent differences between the various theories of personality, the discrepancies between the theoretical viewpoints actually revolve around only four basic issues: the usefulness of the self concept in changing behavior, the importance of the self concept, the amount of change that is possible, and the complexity of the process of change. It is seen that, in addition to being read as single units, the various theories may be meaningfully compared and contrasted on the four basic issues. The importance of this article as a teaching and review aid is stressed.

REFERENCES

1. HALL, C. S., & LINDZEY, G. Theories of Personality. New York: Wiley, 1957. Pp. 548.
2. TRYON, R. C. Cumulative communality cluster analysis. *Educ. & Psychol. Meas.*, 1958, **18**, 3-35.

2120 Roosevelt Street #7
Berkeley, California 94703

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 73-78.

MARYLAND PARENT ATTITUDE SURVEY: A RESEARCH INSTRUMENT WITH SOCIAL DESIRABILITY CONTROLLED* [1] [2] [3]

University of Maryland

DONALD K. PUMROY

A. INTRODUCTION

The area of measuring parents' attitudes toward child rearing is an intriguing one. It seems obvious that the attitudes parents have toward child rearing is related to the way they interact with their children, and this, in turn, should have an effect on the personality of their children. Thus, an instrument for measuring parents' attitudes toward child rearing would be desired for exploring this supposedly obvious relationship. To attest to the popularity of such an instrument, one should note the research done with Shoben's (9) Parental Attitude Scale (PAS) and more recently, and to a greater extent, the instrument developed by Schaefer and Bell (8), the Parental Attitude Research Instrument (PARI). So much research was completed with the PARI that Becker and Krug (1) were able to review the research done and to critique the PARI in the short span of approximately seven years. One difficulty with the PARI pointed out by Becker and Krug is that the response set of subjects is not controlled. Consequently it appeared that true attitudes of the subject were not revealed. The author, on the basis of previous research (6), made a similar observation using the Shoben scale, and over a period of time developed a scale for measuring child-rearing attitudes with one aspect of response set controlled. This instrument is named the Maryland Parent Attitude Survey (hereafter referred to as the MPAS). The MPAS controls for social desirability in a manner similar to the method used in the Edwards Personal Preference Schedule (4). Following its development the MPAS was administered to various groups, norms were

---

* Received in the Editorial Office, Provincetown, Massachusetts, on June 30, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was supported in part by the University of Maryland General Research Board.

[2] A portion of this paper was presented at the American Psychological Association Meeting, Cincinnati, Ohio, September, 1959.

[3] The computer time for this project was made available by the Computer Science Center of the University of Maryland.

established, and *T* scores were calculated. Reliability and interscale correlations were also established. Other studies relating the MPAS to age, sex, and other variables have been conducted. The purpose of this paper is to present briefly the development of the MPAS so that other researchers will know of its availability.[4] It should be noted that the MPAS is strictly a research instrument and should not be viewed as a clinical tool. The author's position is exactly the same as that stated by Bell and Schaefer (2) regarding the PARI.

B. Procedure

The first step in developing the MPAS consisted of accumulating items used in other parent attitude scales and in the preparation of other items from material on child rearing. The items were written in the third person, as was suggested by Shoben (9). (The recent Becker and Krug (1) article points out that the assumption that third person items yield more truthful answers than do first person items has not been borne out.) These items were then sorted so that there would be no duplication and rephrased so they would be consistent with each other. The items were also modified so that they did not refer specifically to mother or father and, thus, would be more flexible for future research. This pool of items was then given to a group of subjects (55 high school seniors, 60 college students, and 63 parents who attended a PTA meeting); there were 96 females and 82 males in the group. The subjects were instructed to answer each item "strongly agree" (SA), "agree" (A), "disagree" (D) or "strongly disagree" (SD), as they thought a *good* parent would. The next step consisted of tabulating the percentage of people who answered each of the four categories for each of the items. Thus, for each item, there was a certain percentage of answers for each category: e.g., item 20 might have been responded to as follows: SA, 50 per cent; A, 50 per cent; D, 0 per cent; and SD, 0 per cent.

Then, the entire pool of items was given to a group of nine psychologists with instructions to categorize each item according to the type of parent the item represented. The type of parent and the description were developed from the literature on parental types. There were originally five categories.

1. *Indulgent Parents*

These parents are child centered; the child is allowed to have his own way in all matters. The child is showered with warmth and affection. While there are attempts at discipline, the child knows the rules can be circumvented. The child is not encouraged to show any initiative, and seldom does he have any responsibilities around the house. Frequently, but for no particular reason other than an impulse on the part of the parent, the child is given gifts and treats.

---

[4] A copy of the MPAS, scoring keys, and *T* scores are available from the author for one dollar. Permission will be granted for reproducing the MPAS.

2. *Indifferent Parents*

These parents have no strong feelings for the child, either positive or negative. These parents show little interest in their children's thoughts or activities; they are primarily concerned with their own activities and prefer not to be bothered by their children. The child has no restrictions placed on his behavior, and is free to do as he pleases.

3. *Disciplinarian Parents*

These parents need and expect fairly strict obedience from the child. The child knows that if he does not comply he will be punished, as the rules are explicitly stated by the parent. This punishment is carried out in a fair and consistent manner. This parent is constantly pushing the child to achieve beyond his ability, forcing him to grow up early.

4. *Rejecting Parents*

These parents are openly and actively hostile toward their children. This hostility is frequently reflected in discipline and punishment. This discipline and punishment seems to be based more on the general negative feelings of the parent than on the behavior of the child. Because of the hostility engendered in the child, these parents frequently feel that children are incorrigible.

5. *Protective Parents*

Protective parents are primarily concerned with seeing to it that the child takes a minimum amount of risks. Consequently, the parents are overly watchful of the child and always alert to possible dangerous aspects of all situations. These parents perform tasks for the child long beyond the time the child is capable of doing the task for himself. The child is not *allowed* to grow up and do things for himself (e.g., feeding, bathing, going to school alone, etc.) for fear that something will happen to him.

C. RESULTS

If six of the nine psychologists agreed as to which category the item belonged, it was retained; if not, the item was rejected. It was found that the psychologists were unable to discriminate between the Indifferent and Rejecting items; and, consequently, these items were combined into the Rejecting category. Thus, there were four scales: Disciplinarian, Indulgent, Protective, and Rejecting. For each item, then, there was an item analysis showing the percentage of subjects who answered the item at each of the four levels of social desirability, and the category the majority of psychologists placed it in as a type of parent. An item was paired with another item if (*a*) it represented a different type of parent, and (*b*) approximately the same percentage of answers fell in each category when answered as a "good" parent should—i.e., in the example above, item 20 would be paired with another item that had been answered: SA, 50 per cent; A, 50 per cent; D, 0 per cent; and SD, 0 per cent. Ninety pairs of items were thus formed with 45 items representing each type of parent. These 90 pairs of items, plus the five buffer items at the beginning, compose the MPAS. The task of the subject is to read the pair of items and decide which of the two most represents his attitude. The test is then scored by determining the number of items in each of the four categories that were chosen. The means and standard deviations of a sample of males (mean age 23.9, *SD* 5.90) and females (mean age 25.28, *SD* 8.63) are presented in Table 1. It would appear from the standard deviations obtained that there is fairly wide dispersion in the attitudes measured.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF MALES AND FEMALES ON THE MPAS SCALES

Scale	Mean	*SD*
Disciplinarian		
Males[a]	25.55	7.10
Females[b]	23.96	6.35
Indulgent		
Males	20.52	6.78
Females	23.05	6.60
Protective		
Males	24.24	6.53
Females	23.99	4.90
Rejecting		
Males	18.50	5.91
Females	18.59	6.04

[a] Total number of males = 196.
[b] Total number of females = 187.

The reliability of the MPAS scales has been measured in two ways, split-half and test-retest (see Table 2). The subjects in the test-retest were college students (males = 30, females = 24), and approximately three months elapsed between the two administrations. In the split-half reliability (Spear-

TABLE 2
RELIABILITY OF THE MPAS SCALES

Scale	Test-Retest	Split-Half	
		Males	Females
	$(N = 54)$	$(N = 45)$	$(N = 45)$
Disciplinarian	.622	.736	.755
Indulgent	.730	.675	.843
Protective	.686	.758	.732
Rejecting	.651	.743	.666

man-Brown), the subjects were college students (males = 45, females = 45). The reliabilities for the four scales vary from .622 to .843, which is similar to other instruments of this nature.

The intercorrelations between the scales were computed even though the data does not satisfy the statistical assumptions underlying correlation: i.e., the measures are not independent. Because the measures are interrelated the correlation should be spuriously high. The correlations are presented in Table 3 and are based on the scores of 54 college students (males = 30, females = 24). There is a negative relationship between the D and I scales and also between the P and R scales.

D. DISCUSSION

The validity of such an instrument is somewhat difficult to establish. Each item, of course, has face validity by the nature in which the item was originally selected. Since the development of the MPAS, several studies have been completed in which the scales have been related to other instruments or behavior: (*a*) As would be expected males score higher on the D scale than do females, while females score higher on the I scale than do the males. (*b*) Age seems to be related to the D and R scales. The older subjects, both male and females, score lower on the D scale and higher on the R scale than do younger subjects. (*c*) The relationship between Edwards' SD scale (5) from the MMPI is not related to the four scales of the MPAS. The correlations range from —.17 to .19, indicating that scales appear to be relatively free of social desirability. (*d*) Brody (3) conducted a study that related mothers' MPAS scores and mothers' rating of their interaction with their children.

TABLE 3
INTERCORRELATIONS BETWEEN THE MPAS SCALES $(N = 54)$

Scale	Correlation		
Disciplinarian	—.694	—.015	.055
Indulgent		—.209	—.394
Protective			—.558

She found that mothers high on the D scales showed more directing and restricting behavior than those that scored low. Also forbidding behavior on the part of the mother was positively related to the mother's score on the R scale. (*e*) The author (7) found that parents in a parent discussion group who attended fewer meetings scored higher on the R scale than those who attended more meetings. He also found that the more meetings the parents attended, the higher they scored on the D scale.

Because of the suggestion of Becker and Krug regarding the control of response set and the studies that have related the MPAS to other measures, it is felt and hoped that the MPAS will make a contribution to knowledge of parental attitudes toward child rearing.

E. SUMMARY

The purpose of this paper was to present a research instrument (Maryland Parent Attitude Survey) designed to measure parent attitudes toward child rearing with social desirability controlled. The rationale, development, reliability of the instrument, and validity studies using the MPAS are presented.

REFERENCES

1. BECKER, W. C., & KRUG, R. S. The Parent Attitude Research Instrument—A research review. *Child Devel.*, 1965, **36** (2), 329-365.
2. BELL, R. Q., & SCHAEFER, E. S. Informal notes on the use of PARI. Unpublished mimeographed report, National Institutes of Health, Bethesda, Maryland, 1959.
3. BRODY, G. F. The relationship between maternal attitudes and mother-child interaction. Paper read at Eastern Psychological Association, Philadelphia, Pennsylvania, April, 1964.
4. EDWARDS, A. L. Manual for the Edwards Personal Preference Schedule. New York: Psychol. Corp., 1954.
5. ————. The Social Desirability Variable in Personality Assessment and Research. New York: Dryden Press, 1957.
6. PUMROY, D. K. The effects of intermittent reinforcement on resistance to extinction and emotional behavior in preschool children. Unpublished doctoral dissertation, University of Washington, Seattle, Washington, 1954.
7. ————. A new approach to treating parent-child problems. Research Grant Report NIH (MH 10029-01), University of Maryland, College Park, Maryland, 1964.
8. SCHAEFER, E. S., & BELL, R. Q. Development of a parental attitude research instrument. *Child Devel.*, 1958, **29**, 339-361.
9. SHOBEN, J. R., JR. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monog.*, 1949, **39**, 101-148.

University Counseling Center
Shoemaker Building
University of Maryland
College Park, Maryland 20740

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 79-90.

MEMORY FOR VERBAL SEQUENCES AS A FUNCTION OF THEIR SYNTACTICAL STRUCTURE AND THE AGE OF THE RECALLING CHILD*[1]

Biometrics Research, New York; and Polytechnic Institute of Brooklyn

SUZANNE SALZINGER, KURT SALZINGER, AND SALLY HOBSON

A. INTRODUCTION

One of the difficulties in assessing the development of syntax in the language of young children lies in the discrepancy between the structure of the utterances the child normally emits and the structure of the utterances of other speakers to which the child is able to respond. Children make responses to structured verbal sequences that they can not or do not ordinarily emit on their own (7). The present study concerns itself with an examination of the preschool child's ability to respond differentially to verbal sequences varying in syntactical structure.

Some of the more common structured word sequences—e.g., simple declarative sentences like "I see you."—may be first perceived as single units due to the fact that one of the major formal separations characteristic of written language—namely, separation between words—is not consistently made explicit in speech and is therefore not available to the nonreading young child. Kagan (5) has suggested that inflection and loudness function as the stimulus characteristics that make these units perceptually salient in the child's verbal environment. Such units are acquired early and, due to their high frequency of occurrence in English, become very strong responses.

As the child is exposed to more adult speech, he begins to perceive the various parts of these gross single units as units themselves. The process by which this comes about is that of discrimination: shorter phrases, or words, or morphemes are heard in different verbal contexts and in this way acquire a unitary aspect of their own. Evidence for the effect of the discriminative properties of bound morphemes has been given in experiments using nonsense syllables as basic word stems (1, 2, 11). Given such a constantly increasing

---

* Received in the Editorial Office, Provincetown, Massachusetts, on July 5, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The work described in this study was supported by a grant No. MH07477 from the National Institute of Mental Health. Thanks are due to Dr. Joseph Zubin for his interest and advice.

repertoire of discriminated units, we find that single words become the primary unitary form. In time the child can handle, in addition to the refined small units, such as words, all the earlier "units," such as simple sentences, now in the form of syntactically linked sequences.

Assuming that children are exposed to a corpus of English that represents more general English speech, the extent to which the various units cohere for the child should gradually tend to approximate the frequency with which such units are found in English. More specifically, if we measure the strength of these units by the ability of the child to recall word sequences ordered in various ways, the extent of the increase in recall for a given word sequence should depend upon how often its parts are heard in sequence or as single units, and how often they are broken up in English speech.

The present study is designed to examine the strength of cohesion of these syntactical units for children between the ages of 3 and 6 years. Making the assumption that the linking of words through the discriminative properties that constitute their syntactical structure reduces the number of units that need to be encoded in an immediate memory situation (9, 10), we have used a recall measure as an index of the strength of the syntactical structure, and reorderings of a simple declarative sentence as a way of varying the degree of syntactical structure in the stimulus material. More specifically, the declarative structure was presented intact and with mutilations occurring in less and less likely positions: *viz.,* subject—predicate, verb—object, noun—modifiers, and complete reversal of word order. If the extent of the children's recall is positively related to the relative frequency characteristics of the various word sequences, we can conclude that the children have in fact made the discriminations necessary for utilizing the structures in encoding. If we find no order of recall consistent with the relative frequency characteristics of English word orders, we can conclude that the children have not yet made the necessary discriminations. Such children—i.e., very young children—might still be functioning at a stage where only the gross initial units can be utilized to improve recall; other subdivisions of these units can not yet be discriminated by them and therefore can not increase their ability to recall such sequences over word sequences arranged in random order.

The study, furthermore, was set up to examine both short (four-word) sequences as well as long (six-word) sequences with the purpose of replicating the effect of disruption of syntactical structure on sequences that tax the memory span in this age range to different degrees.

B. METHOD

1. *Subjects*

The *S*s consisted of 49 children, ranging in age from 3 to 6 years, who attended a private day nursery and kindergarten.[2] The group can be characterized as middle class with most of the children coming from professional and semiprofessional homes.

2. *Apparatus*

The apparatus consisted of a table upon which was placed a large burlap screen to separate the *S* from the *E*. In front of the screen, on *S*'s side of the table, was a *papier-maché* clown mounted in a colorful box; his nose was made of a 7½-watt red light bulb which could be turned on and off by *E*. The same apparatus was used previously in a study on operant conditioning of speech (12). A transparent plastic chute, through which small trinkets could be delivered, was imbedded in the screen and fed into a small clear plastic box on the table next to the clown. A wooden toy mallet was also placed on the table in front of *S*. Two tape recorders were placed on *E*'s side of the screen, one for presenting to *S* the stimulus word sequences that were recorded in the clown's voice, and the other for recording *S*'s attempted imitation of the word sequences. A foot pedal for controlling the presentation of stimulus sequences and a microphone which was hung around *S*'s neck completed the apparatus. The entire testing area of the room was enclosed with striped drapes resembling a circus tent.

3. *Procedure*

Before testing the *S*s individually, the two *E*s visited them in their classrooms and read them an illustrated story about the adventures of "Happy the Clown." The story not only served the purpose of introducing the clown but also elicited from the *S*s, in a familiar setting, the types of responses later required during individual testing.

Each *S* was then brought individually into the experimental room which, he was told, was "Happy's tent." *S* was seated in front of the *papier-maché* clown at the table and the microphone was placed around his neck. The two

---

[2] Stephen Wise Free Synagogue Nursery School and Kindergarten. The authors would like to acknowledge the excellent cooperation they received from both the school and the synagogue, and particularly express their appreciation to Mrs. Regina Gilbert, the school director, and her staff and Rabbi Edward Klein.

*E*s showed him the clown, the trinket box and chute, and the mallet, and then told him that the clown would speak to him and play games with him.

The clown then partly sang, partly spoke the following tape-recorded instructions (S^r indicates reinforcement):

> (Whistling). I'm so happy (S^r_1 = clown's nose lighting up), I'm so happy (S^r_1), I'm so ha-pa-pa-ha-pa-pa happy (S^r_1). Whoops, well hi there. My name is Happy. What's yours? (*S*'s response) (S^r_1). That's a nice name. Say, how do you like that? Whenever I'm happy my nose lights up. You know what? I like you. If you say "hello" to me I'll give you a present. (*S*'s response) (S^r_1) (S^r_2 = delivery of trinket). Do you remember how we said "boo" to the lion? Well, if you say "boo" for me now, I'll give you another present. (*S*'s response) (S^r_1) (S^r_2).

The clown then instructed *S* to imitate two identical series of stimuli, one of rhythmic "boos" and one of rhythmic mallet taps.[3] Every response was followed by S^r_1 and S^r_2. The following instructions were then given by the clown for the verbal response task:

> Great! You're just great. You're just wonderful. Say, do you know my favorite game? (*S*'s response) (S^r_1). Well, I call it silly mixed-up words. The silly mixed-up words I say are so silly and so ridiculous that even ferocious lions laugh at me. Say, I'll bet that you can say some of my favorite mixed-up words. O.K.? (*S*'s response) (S^r_1). Now listen carefully, and after you hear me say my silly thing then you say it right after me. O.K.? (*S*'s response) (S^r_1).

Twelve sequences of words were then presented, one at a time, for *S* to imitate. If *S*'s response was not correct, the word sequence was repeated once for 3-year-old *S*s and, if necessary, twice for older *S*s before S^r_1 and S^r_2 were delivered. Reinforcement was withheld only if *S* refused to attempt imitation. Analysis of these data was based only on the first attempted imitation. Five different types of word sequences were used: (*a*) simple declarative sentence, (*b*) reversal of subject and predicate, (*c*) reversal of verb and object, (*d*) reversal of noun and modifier, and (*e*) complete reversal of declarative sentence.[4] Two sequence lengths, one four words and the other six words, were used to sample each type of reordered sequence. The following is one of the four lists used in the experiment:

---

[3] The two rhythmic tasks served to accustom *S* to responding to the clown prior to the verbal response part of the session, as well as to give some indication of the number of nonverbal units the children could process.

[4] The intonation and rhythmic patterns adopted for each stimulus were based on standard English intonation for the intact parts of each stimulus sequence, with pauses between the intact portion of about the duration of the pauses that ordinarily separate phrases.

Practice 1 —The fox plays games.
Practice 2 —The black wolf likes short coats.
3a—The horse pants wears.
3b—The young pig new hats buys.
1b—The old cow chews pink gum.
1a—The dog reads books.
5a—Eggs eats mouse the.
5b—Pants tight wears horse fat the.
4b—Dog small the reads books good.
4a—Cow the chews gum.
2a—Buys hats the pig.
2b—Eats cold eggs the white mouse.

The first two sequences were identical on all four lists. A schematic representation of the makeup of the four lists is shown in Table 1. Lists I and III and lists II and IV are identical in content. Lists I and II and lists III and IV are identical in terms of the order of presentation of the different types of word sequences. Lists I and IV and lists II and III are identical in terms of the order of presentation of four- and six-word sequences.

TABLE 1
THE ORDERING OF THE STIMULUS MATERIAL

Content A		Content B	
List I	List III	List II	List IV
3a	4b	3b	4a
3b	4a	3a	4b
1b	2a	1a	2b
1a	2b	1b	2a
5a	1b	5b	1a
5b	1a	5a	1b
4b	5a	4a	5b
4a	5b	4b	5a
2a	3b	2b	3a
2b	3a	2a	3b

Note: The numbers 1, 2, 3, 4, and 5 refer to the different types of word orders; and a and b refer to four- and six-word sequences, respectively.

Each *S* was tested twice, the time interval between the test and retest being approximately three months. *S*s were given lists at retest that differed in content from their test list but in which the order of types was the same. Most of the *S*s were retested in the same order in which they were tested, but because of absences this was not always possible. Forty-nine *S*s were tested initially and

45 of the *S*s were retested. The remaining four *S*s were not available for retesting for reasons unrelated to the experiment.

As a comparison group, seven college students were presented with one of the tape-recorded stimulus lists and asked to write down their responses.

C. RESULTS

Five age groups of nine children each were formed of the 45 children who had been given both the test and retest. To achieve the groupings, the children were ordered according to their age at the time of initial testing and successive groups of nine were taken to constitute each group. The mean ages and age ranges for each group can be found in Table 2.

TABLE 2
MEAN AGE AND AGE RANGE IN YEARS FOR EACH OF FIVE AGE GROUPS

Age group	Test		Retest	
	M	Age range	*M*	Age range
I	3.2	3.0—3.3	3.5	3.3—3.6
II	3.6	3.4—3.7	3.9	3.7—4.0
III	4.0	3.8—4.4	4.3	4.1—4.6
IV	4.5	4.4—4.7	4.8	4.7—5.0
V*	5.0	4.8—5.8	5.3	5.1—6.1
Total	4.1	3.0—5.8	4.4	3.3—6.1

* One of the children in this group was considerably older than the others.

Each child's recall of a given stimulus word sequence was scored in two ways: (*a*) the largest number of correct words recalled in the order presented—the longest sequence recalled, and (*b*) the number of correct words recalled independent of order. Analyses of variance using a three-way layout in a mixed model (13) were applied separately to the two types of scores (mean longest sequence recalled and mean number of correct words recalled) at both test and retest for the four- and six-word sequences in order to determine the effect of age and structure of sequences on recall.

Figure 1 presents the means of the longest sequence of words recalled as a function of age with syntactical structure represented as a parameter.

Significant differences among the age groups were found ($p < .05$). Inspection of Figure 1 shows that the increase in age is positively related to an increase in recall scores for the four- and six-word sequences at test and retest. This result holds as well for data scored as mean number of correct words recalled ($p < .01$).

Differences in recall of the five types of word sequences when scored as the mean longest sequence correct (see Figure 2) were found to be statistically significant for both the four- and six-word sequences at both test and retest ($p < .005$). The data show that immediate memory decreases with reorderings of a simple declarative structure in such a way that best recall is found for the intact declarative sentence (1), next best for the subject and predicate reversal (2), next for the reversal of verb and object (3), next for the reversal of nouns and their modifiers (4), and poorest for the complete sentence re-

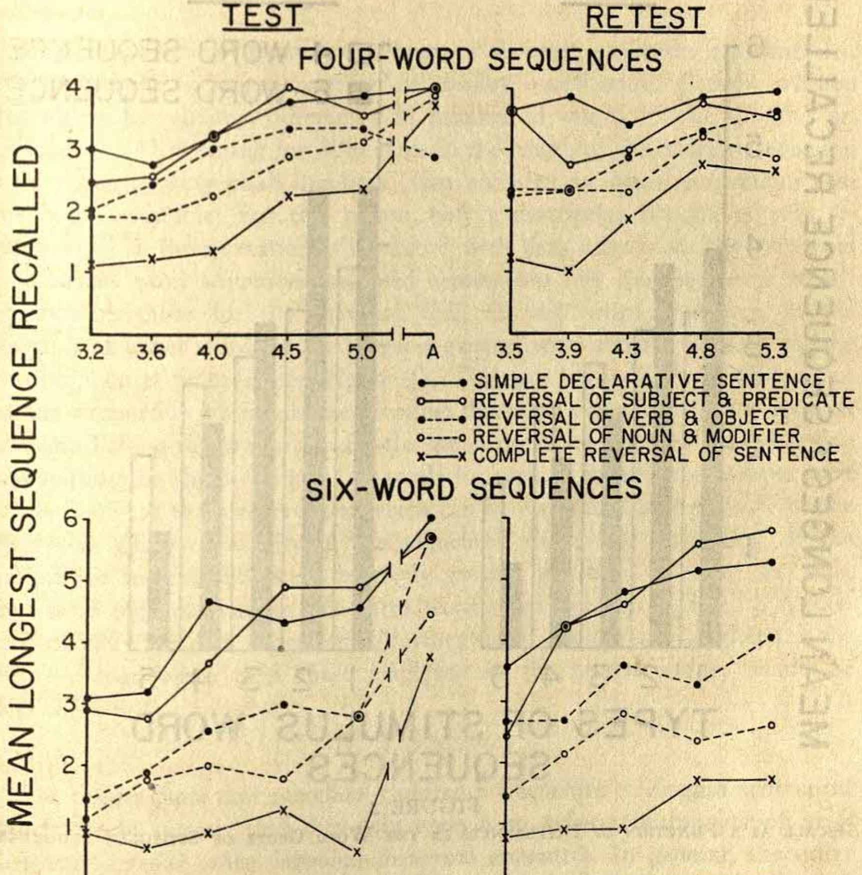

FIGURE 1

RECALL AS A FUNCTION OF AGE WITH SYNTACTICAL STRUCTURE VARIED AS A PARAMETER versal (5). The data represented by mean number of words recalled yielded the same results ($p < .005$).

The interaction between age and recall of the different sequences represented by the mean longest sequence correct as tested by the analysis of variance indicates no statistically significant effect ($p > .05$). Figure 1 shows that memory

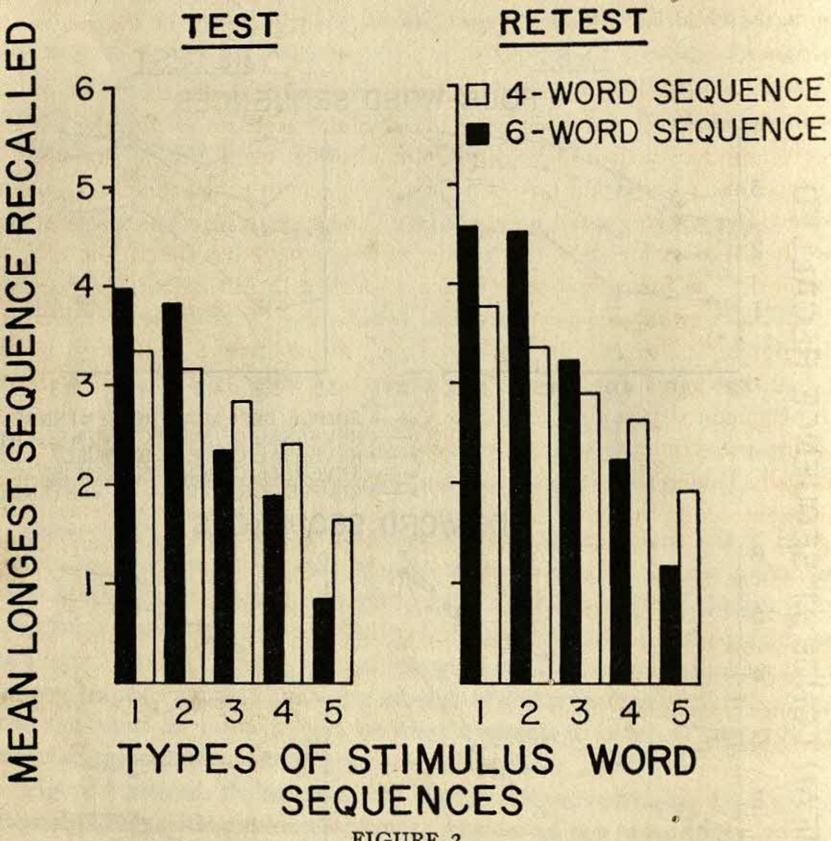

FIGURE 2

Recall as a Function of Differences in the Word Order of Stimulus Sequences

for the different word sequences lines up at each age essentially as it does for the entire group. This result holds for retest as well as test despite a general rise in memory for most word sequences at most ages.

The result holds for recall scores of number of words correct ($p > .05$) except for the four-word retest condition ($p < .025$), which is most probably due to a ceiling artifact produced by a combination of conditions that makes the task too easy for the children.

Comparable data on the group of seven adults showed generally the same relative ordering among recall scores of the different types of word sequences. The mean longest sequence recalled for adults is plotted as a discontinuous point on the graphs for initial test on Figure 1 for purposes of comparison with the children, although these data were not included in the analysis of variance.

Any comparison of memory for four- or six-word sequences confronts one with the choice of (*a*) comparing the absolute recall scores directly, without correcting for obvious differences in number of words in the stimulus sequences or (*b*) equating for differences in the length of the stimulus sequences by dividing by very small numbers (four and six) to obtain proportions that are highly unstable. For this reason, only a descriptive comparison was attempted. The interpretation is in accord with both tabulations. As the types of stimulus word sequences presented become less like English syntax, recall deteriorates more for the six-word than the four-word sequences, to the extent that fewer words of the longer sequences were actually recalled for the less common structures (see Figure 2). This was borne out by consideration of the proportion of words constituting the mean longest sequence recalled as well. Taking the proportion for the intact declarative sentence as a baseline and subtracting the proportions of recall for each of the other sequences from it, we find that these differences, which can be considered as a measure of the disruptive effect of modifying the sequences, show that reordering of the declarative sentence produces relatively greater deficits in recall of six- than four-word sequences as the structures become less common to English syntax. In both cases the data suggest that further study, using materials that tax the memory span more than those employed in the present study, would be advisable.

D. DISCUSSION

The results show that preschool children are sensitive to English syntactical structure and suggest that they rely upon it to reduce the number of units they must encode when listening to verbal sequences. In general, the order in which the modification of syntactical structure reduced recall remained constant throughout the age range studied and probably reflects the relative frequency of occurrence with which these structures are found as units or sequences in English. It appears, therefore, that by three years of age children have already developed responses to speech that reflect these general linguistic frequency characteristics. Further development shows a general increase in the ability to recall verbal sequences but maintains the relative order of response to speech characterized by various syntactical structures. The same ordering was also reflected in the performance of adult subjects, particularly where six-word sequences—which tax the memory span—were used. A similar lack of interaction of grammatical forms with age was also reported in a study by Martin and Jones (8), who found no difference between children and college students.

Furthermore, in a study by Karpova (6) as cited in (3), which tested 5 to 7 year olds' awareness of the units in sentences, the ordering of units to which the children responded coincided well with the ordering of units that aided in recall in the present study. Karpova found first that children tended to hear sentences in two parts: namely, a subject and a predicate. The children in this study recalled this reordered sequence almost as well as the intact declarative sentence. When more than two parts were heard, they were most often concrete nouns, then verbs and adjectives, and least often function words. The children had great difficulty in recalling reversed noun-modifier pairs where the modifier always included the function word "the."

The fact that breaking up a declarative sentence into its subject and predicate produces little observable difference from recall of the intact sentence, and Karpova's finding that children hear sentences as a separate subject and predicate, suggests that these phrases may be the basic perceptual units for encoding simple declarative sentences. Martin and Jones (8) also found a similar type of equivalence between single words and certain phrases in an associative learning task.

That certain syntactically linked word sequences do function as units was indicated by Huttenlocher (4), who found in a study of 4½- to 5-year-old children that being able to separate word pairs was a critical factor in the ability to reverse members of the same pairs. As a result, members of grammatical pairs were found to be more difficult to reverse than nongrammatical pairs. If syntactical structure does function in such a way as to produce units, then words that occur together most frequently should be most difficult to recall if they are reordered. Such seems to be the case in the present study.

There is some indication, although not conclusive, in the present study that the syntactical structure of a verbal sequence becomes increasingly critical for recalling that sequence as its length increases beyond that of the memory span. It appears that extreme modification of common syntax may even produce an interference effect such that longer sequences, which when syntactically structured enable *S*s to recall a great deal of information, come to be less well recalled than are shorter sequences. Such a notion would be consistent with Miller's chunking hypothesis (9) and could be easily tested further in a study that varied more extensively the length of sequences as a parameter of the experiment.

E. SUMMARY

Based on the assumption that the syntactical structure in a sequence of words reduces the number of units that a listener must process, 45 preschool children, aged 3 to 6 years, were tested on their immediate memory for various reorderings of a declarative sentence.

Recall increased with age and similarity of the verbal sequence to the usual structure of a simple declarative sentence. The results were interpreted to demonstrate the early development of the child's sensitivity to linguistic frequency characteristics of English: i.e., the frequency with which parts of speech occur in given orders.

REFERENCES

1. EPSTEIN, W. The influence of syntactical structure on learning. *Amer. J. Psychol.*, 1961, **74**, 80-85.
2. ———. A further study of the influence of syntactical structure on learning. *Amer. J. Psychol.*, 1962, **75**, 121-126.
3. ERVIN, S., & MILLER, W. R. Language development. *62nd Yearbook Nat. Soc. Stud. Educ.*, 1963, **62** (Part I), 108-143.
4. HUTTENLOCHER, J. Children's language: Word-phrase relationship. *Science*, 1964, **3603**, 264-265.
5. KAGAN, J. Discussion of Jenkins and Palermo's mediation processes and the acquisition of linguistic structure. *Monog. Soc. Res. Child Devel.*, 1964, **29**, 169-173.
6. KARPOVA, S. N. Osoznanie slovesnogo sostava rechi rebyonka doshkolnogo vozrasta (Awareness in the word content of speech of a preschool child). *Vop. Psikhol.*, 1955, **I**, 43-55.
7. LENNEBERG, E. H. Understanding language without ability to speak: A case study. *J. Abn. & Soc. Psychol.*, 1962, **65**, 419-425.
8. MARTIN, J. G., & JONES, R. L. Size and structure of grammatical units in paired-associate learning at two age levels. *J. Exper. Psychol.*, 1965, **70**, 407-411.
9. MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81-97.
10. SALZINGER, K. Some problems of response measurement in verbal behavior: The response unit and intraresponse relations. Paper read at the Conference on Methods of Measurement of Change in Human Behavior, Montreal, Canada, 1962.
11. SALZINGER, K., & ECKERMAN, C. Grammar and the recall of chains of verbal responses. Paper read at the Eastern Psychological Association, Atlantic City, New Jersey, 1965.

12. SALZINGER, S., SALZINGER, K., PORTNOY, S., ECKMAN, J., BACON, P. M., DEUTSCH, M., & ZUBIN, J. Operant conditioning of continuous speech in young children. *Child Devel.,* 1962, **33**, 683-695.
13. SCHEFFÉ, H. The Analysis of Variance. New York: Wiley & Sons, 1959.

Biometrics Research
722 West 168 Street
New York, New York 10032

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 91-93.

A PRELIMINARY STUDY ON FACTORS IN LEARNING IN A HOW-TO-STUDY COURSE*[1]

Department of Psychology, New York University

M. E. TRESSELT

A. INTRODUCTION

Whenever experiments are performed using a group outside of the laboratory, numerous intervening variables are operative that may play a more important role in the experimental design than was originally assumed or that might, upon investigation, indicate different lines of research.

In 1951, Tresselt and Richlin (2) published a short report concerning the efficacy of the how-to-study course at New York University. They reported that the averages of the students increased significantly with a p value of less than .01 when the grades before taking the course and the grades after taking the course were compared. They selected three variables—intelligence, personality, and study habits—for emphasis and were able to show that these variables played differential roles. They summarized their findings in the following way. (*a*) The students having a high amount of ability and personality and low study techniques showed the highest amount of improvement. (*b*) The group having a high amount of ability and study habits and a low personality adjustment showed differential improvement. When this last group was subdivided into a "severe" degree of personality maladjustment and a less "severe" degree of maladjustment, the latter showed the greater improvement. (*c*) The group having a high ability and a low personality adjustment and study techniques showed less improvement than the other two categories but essentially resembled the *b* group. (*d*) The other combinations of the variables also showed improvement but the amount of improvement was less, although still statistically significant.

Later Richlin and Tresselt (1) added to their original data confirming the efficiency of the course, but using the biserial r, were able to show that intelligence was by far the most important variable and the study habits the least important, with the importance of the personality factor varying between the two extremes.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on July 19, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to thank Dr. Milton Richlin for his cooperation in this study.

B. Procedure

A questionnaire was devised which contained 29 questions of three types. Some statements were incomplete and the subjects were asked to complete the sentence. Some questions could be answered by "Yes," "No" or "Don't know." The rest of the questions were rating scales. The questionnaire was given ostensibly to find out how the course could be improved and, in partial support of this idea, the students were told not to sign their names (although by having two experimenters in the room it was possible to check what paper was written by which student). The questionnaires were administered three times: one month, two months, and three months after classes started. The questions concerned the subject matter of the course, the relevancy of lectures, the attitude toward the instructor, the importance of interviews, and the reaction of the student to the class and toward his own improvement.

C. Results

The first question (Question 5) raised by the experimenter involved the belief of the students that this course was going to help them. Initially there were many more students who expected to be helped than actually were helped by the lectures and interviews, but in two months this expectation and achievement became increasingly similar: i.e., 32 expected to be helped and 19 felt as if they were helped in the first month, as compared to 31 who expected and 28 who felt they were helped by the third month. These data warrant further analysis.

On the first administration of the questionnaire, there were 13 students who answered "Yes" to the questions of whether they expected help and felt that they had been helped. The average *IQ* (Wechsler-Bellevue) of these subjects was 115.5. The 10 additional subjects who answered "Yes" to both questions on the third administration had an *IQ* of 109.3. It seems, then, as if those subjects with the higher abilities recognized the expectation and actual aid sooner than did subjects of lesser intelligence.

In grade average, the 11 students (two of the original 13 had no classes prior to taking the course) had an average of 71.0 before taking the course and 77.6 during the course; those responding on the third enquiry had an average of 71.4 before the course and 77.0 during the course. It is also of interest to find that the two who answered "Don't know" to Question 5 in the first questionnaire had an average of 71.4 before and 69.3 during the course. Although none of these statistics is significant in itself, all the statistics seem to suggest the hypothesis that those who believe they will be helped by the course will be helped to a greater extent than those who are doubtful of the benefits of the course.

To some extent it may be argued that the belief in aid, resulting from the course and the actual change in grades, is really a result of the change in the level of aspiration. In order to test the relationship of level of aspiration to other known factors, the subjects were asked what grades they expected to get over the period of three months. A correlation between *IQ* and goal discrepancy was —.13. When the students were divided into two groups either having the intelligence to complete college work satisfactorily—i.e., an *IQ* of 115 or more—or having a negative amount of the intelligence factor—i.e., 114 or less—the high *IQ* group had a discrepancy average of 3.4 grade points, while the low *IQ* group had a larger discrepancy score of 6.3—the low *IQ* group estimated their aspiration level higher than their achievement level.

If there is a difference in the level of aspiration and intelligent students show a greater improvement with the course than do less intelligent students, the level of aspiration should reflect the possible improvement in college grades. The rho between goal discrepancy and grade improvement was +.55, indicating that the less the goal discrepancy, the greater the possibility of grade improvement.

D. Summary

In order to study the lesser variables in aiding students in a how-to-study course, an "anonymous" questionnaire was administered three separate times during a semester to a total of 35 subjects. The results seem to suggest that (*a*) those who believed they would be helped by the course were helped to a greater extent than those who were doubtful, and (*b*) the more intelligent students seemed to recognize the possibility of help earlier in the course than did the less intelligent students.

References

1. Richlin, M., & Tresselt, M. E. Further research on the differential prognosis in a college study methods course. *Amer. Psychol.,* 1951, **6**, 368-369.
2. Tresselt, M. E., & Richlin, M. Differential prognosis in a college study methods course. *J. of Psychol.,* 1951, **31**, 81-89.

Department of Psychology
New York University
Washington Square
New York 3, New York

of the course will be helped to a greater extent than those who are doubtful of the benefits of the course.

To some extent it may be argued that the belief in aid, resulting from the course and the actual change in grades, is really a result of the change in the level of aspiration. In order to test the relationship of level of aspiration to achievement targets, the subjects were asked what grades they expected to get after the period of three months. A correlation between IQ and goal discrepancy was —.13. When the students were divided into two groups either having the intelligence to complete college work satisfactorily—i.e., an IQ of 115 or more—or having a negative amount of the intelligence factor—i.e., [illegible] or less—the high IQ group had a discrepancy average of 1.4 grade points, while the low IQ group had a larger discrepancy score of 6.3—the low IQ group estimated their aspiration level higher than their achievement levels.

If there is a difference in the level of aspiration and intelligent students show a greater improvement with the course than do less intelligent students, the level of aspiration would reflect the possible improvement in college grades. The rho between goal discrepancy and grade improvement was —.35 indicating that the less the goal discrepancy, the greater the probability of grade improvement.

D. Summary

In order to study the belief variable in aiding students in a how-to-study course, an "anonymous" questionnaire was administered three separate times during a semester to a total of 35 subjects. The results seem to suggest that (1) those who believed they would be helped by the course were helped to a greater extent than those who were doubtful, and (2) the more intelligent students seemed to recognize the possibility of help earlier in the course than did the less intelligent students.

References

1. Kornreich, M., & Kramer, M. E. Further research on the differential effectiveness of a college study methods course. J. educ. Psychol., 1951, [illegible].
2. Kramer, M. E., & Kornreich, M. Differential progress in a college study methods course. J. educ. Psychol., 1951, 42, [illegible].

Department of Psychology
New York University
Washington Square
New York 3, New York

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 95-100.

COMMUNICATION AND MEANING*

Campbell-Ewald Company, Detroit

DANIEL K. STEWART

A. INTRODUCTION

Human communication is concerned with the elicitation of intended meaning. The study of communication, therefore, must include an understanding of the meaning of "meaning." However, a problem in regard to the analysis of "meaning" is to select a symbol system, a language system, that will point toward the direction of significant research on this subject. The analysis of "meaning" requires a language system that will not do violence to the psychological phenomena of which it is concerned. Symbolic order is necessary for "meaning" anytime, but particularly in its formal analysis (5). The study of communication is in immediate need of a psychology valid for its subject matter and a philosophy explanatory of the human mind (6).

The psychological basis of the analysis of "meaning" to be presented in this paper has been succinctly described by Blanshard as follows: "to mean is a process of a conscious mind. Arrangements in nature do not mean. Sentences by themselves never mean. Nothing actively means but minds, and though minds may make use of sentences or kings or churches as aids in their thinking, their thoughts are certainly not reducible to the machinery that they use" (1, p. 153).

Now, at the outset, we observe that there are several factors operating in the communication process. The use of *language* is one of them. *Judgment* and *emotion* are others. Important as these factors are, however, it is not the author's intention here to show the nature of their intervention in the communication process. It is noted, briefly, that language should not be confused with the meaning it is employed to elicit. In the interpretation of communicative discourse, if we exercise our judgment at all, we choose between two or more discrete ideas. If we become emotional in response to discourse, we do so within the framework of specific ideas.

The point is that *meaning* is but one part of a complex whole, and that this particular factor would seem to be *the* basic psychological element to

---

* Received in the Editorial Office, Provincetown, Massachusetts, on July 11, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

which all the other factors are subservient. Its existence is necessarily prerequisite to the existence of the others. Deny it, and *mind* itself is denied.

Therefore, the fruitfulness of any experimental design constructed for the purpose of measuring the effects of communication is seen to demand evidence—above all else—either for or against the proposition that the speaker's intended meaning is, indeed, that very meaning elicited by the interpreter. To speak plainly, this is the acid test of any type of conscious communication and therefore, quite properly, the primary responsibility of communication research. If research is not brought to bear on the very objective of communication—the elicitation of intended meaning—it becomes exceedingly difficult to justify all of the contingencies associated with it.

B. An Hypothesis on the Meaning of "Meaning"

Discourse on the nature of meaning is really discourse on the nature of reality, truth, or being—whichever synonym is preferred—and the search is an inductive or dialectical procedure. It is going to the first principles: it is a philosophical enterprise. In this sense, W. J. Oates says that "communication involves philosophy, that it is ultimately a philosophical enterprise . . ." (3, p. 36), and Werkmeister asserts that "without this venture into ontology the problem of communication, and therefore of meaning, cannot be solved completely" (6, p. 266).

Thus, the task is not one of deductive or logistic argument because the meaning of the propositions of such arguments depend upon what one has already assumed the nature of meaning to be. Such propositions are asserted in the argument because they are *assumed* as being meaningful. And, thus, if the proposition to be arrived at is one asserting something about the meaning of "meaning," any deductive argument would be circular because its propositions would require this in the establishment of its conclusion. In other words, the *meaning* of the premises, as determined by an unstated antecedent condition, cannot be assumed where the nature of *meaning* is the declaration in the conclusion to be proved.

Therefore, we cannot make this assumption on the nature of meaning because it is precisely this nature that is our task to investigate. Once we feel we know something about its nature, then, upon assuming this extralogically, we may present valid explanations on its manifestations.

Nor is it the case that there can be a multitude of meanings of "meaning" that are all equally correct. There can be only one true explanation—i.e., one that is reflective of reality—and it is our task as researchers to find out which one it is. As Aristotle reminds us, "where there is not one meaning (of 'meaning'), there is no meaning."

When an individual perceives a physical object—e.g., a word—we note that on occasion an idea is a consequence of the involved perception. Such an idea can be labeled an Immediate Perceived Thought, IPT (2, p. 3). Thus, the first question is: does any discrete idea possess meaning in itself? And the answer is "no." Ideas with meaning always are associated with one or more other ideas (4, p. 891). In fact, as actual practice demonstrates, any given idea cannot be forced into isolation and at the same time provide meaning.

But if an idea is ontologically discrete from its meaning, then what is meaning?

The hypothesis the author wishes to offer is this: the term "meaning" is a physical symbol complex conventionally employed to label that psychological state of affairs wherein given ideas have been harmoniously juxtaposed by the mind. It is a psychological experience of order.

Relative to the universe of ideas within a given mind, such ideas constitute a nonformal psychic system. This is another way of saying that even as descriptive propositions alone never explain anything, so it is in regard to those discrete ideas of which descriptive propositions are the physical representatives.

This does not imply that the mind is necessarily an unordered collection of ideas. It does imply, however, that the ordering efficacy and existent harmony among ideas is a function of the individual mind, and, therefore, varies from one mind to another. In fact, if the term "intelligence" possesses any significance at all, it is in reference to this differential ability of discrete minds to order their own ideas. Disordered minds and nonordered minds are also relevant to this point.

"Meaning" is a term used to describe that state of mind wherein given ideas are associated or juxtaposed to each other harmoniously. The ontogenesis of meaning is a product, therefore, of a mental operation that generally goes under the label of "judgment." What is involved here is a selection by the mind of certain ideas among a "plurality" of ideas, or, more accurately, a selective attending by the mind of its own reality. "To judge" is the language used to describe an active selection, and "judgment" is the past tense record of such a selection.

Meaning, therefore, when possessing being, is always an ideational complex entailing two or more ideas. It exists as a certain ideational relation. Because of this, a psychological analysis of ideational relations becomes critical to the analysis of the meaning of "meaning," and any subsequent understanding of the communication process. But ideational relations that exist between ideas raise problems that are not easily resolved. On this point Benjamin Whorf remarks:

> I have not been able to find a term that I need to denote a kind of connection or relation, approximation, closeness, allied character, between ideas. The only psychological term I know of that expresses connection between ideas is "association," but this has quite a definite meaning and one that will not do for the meaning I have in mind. The "connection" of ideas, as I call it in the absence of any other term, is quite another thing from the "association" of ideas . . . (8, pp. 35, 36). I might say that my mental image of the relation is not at all one of ideas hitched together by bonds of attachment which they possess like miniature hooks and eyes. It is more a concept of continuity, with the ideas as relative locations in a continuous medium (8, p. 38).

C. Relations and Meaning

In regard to psychological phenomena, the term "relation" is open to equivocation because of its ambiguity. For example, the expressions "ideas are related" and "relations exist" both employ the term "relation," but in different ways. The referent in the first expression is a set of associated or juxtaposed ideas, while the referents in the second expression are such physical complexes as "greater than," "implies," "and," etc., or the ideas of these.

For example, in the statement "Harold is the father of Walter" we note, in the first place, that it is reflective of a set of ideas in association with each other. Secondly, we note the physical existence of a relation denoted by the graphic marks "father of." And, thirdly, we note (*a*) that the graphic marks "father of" are not in the mind, (*b*) that the graphic marks "father" and "of" are not in the mind, but (*c*) that a discrete idea of the physical unit "father of" is in the mind relative to the meaning of the whole sentence. And it is the fact of (*c*) that is quite important in regard to the ideational disposition of relations.

It is apparent that there is something quite fundamental occurring in the mind that determines the quantity of ideas actually elicited from discourse. Moreover, it is noted that this quantity of ideas does not necessarily correspond to the quantity of perceptual stimulations involved. The least that can be asserted from this state of affairs is that if ideas are the immediate consequences of conscious perceptual activity (IPT's), then it must be entailed that the physical phenomena involved in such perceptual activity will have ideational analogues *of some kind and quantity*. Therefore, there will exist some one idea reflective of the relation presented in the example given. While this prosaic conclusion is not all that we would like to know in this rather basic operation of the mind, it is quite important because, whatever we find out about these particular ideas in the future, they are what constitutes the structural or "connective" factors within units of meaning. If we consider "ideas as relative locations in a continuous medium," relational ideas are the "concepts of continuity" that Whorf was concerned about.

Therefore, "relations"—as components of language systems—are physical symbols used in conjunction with other physical symbols to describe psychological or physical states of affairs. Such relations have ideational counterparts. In terms of ideational states of mind, physical symbol complexes (e.g., "Harold is the father of Walter") are employed to explicate a certain juxtapositioning or pattern of ideas. In terms of meaning, some of these ideational juxtapositionings will be harmonious or compatible and some will not.

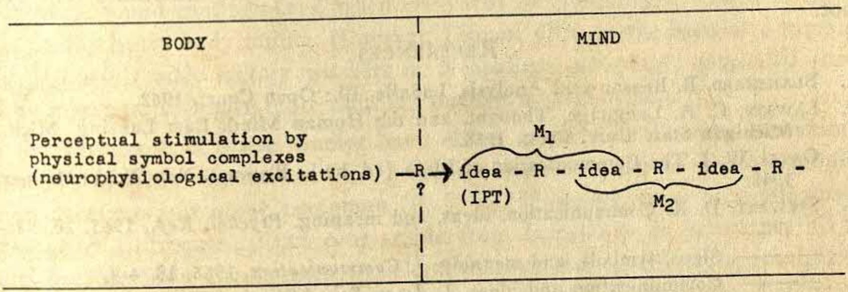
BODY
MIND
Perceptual stimulation by physical symbol complexes (neurophysiological excitations)
R
?
idea - R - idea - R - idea - R -
(IPT)
M1
M2

FIGURE 1
AN ILLUSTRATION OF THE HYPOTHESIS REGARDING THE MEANING OF "MEANING"

Thus, having said this about the term "relation," let us say that "meaning" refers to relational contingencies among ideas; and the meaning of the term "relation," when employed in such an ideational context, refers to that idea wherein the other ideas within the set are structured or patterned in the manner of their meaning as a whole or unit.

Figure 1 illustrates the hypothesis regarding the meaning of "meaning."

In conclusion, "meaning" is analyzed as the experience of order. The "experience" is strictly psychological, and the "order" is strictly ideational. "Meaning," therefore, refers to a specific disposition of mind wherein certain of its ideas exist in harmonious juxtaposition to each other.

Only *ideas* are the objects of meaning. Physical phenomena, such as words, cannot be meaningful *per se*. "Meaning" is a distinctly human property, a *mind* property; and the attribution of this property to physical phenomena can result in research that is irrelevant to the objective of human communication.

Only ideas can provide meaning, and no single idea is meaningful in itself. This conjunction says that whereas ideas are the objects of meaning, meaning is not a property of those ideas. Rather, it is the case that meaning is a function of the mind having ideas as arguments.

In other words, there are no "meaning" ideas *per se*. A given set of ideas does not "view" itself as being in a "meaning" disposition. "Meaning" is a product of two or more ideas existing in harmonious juxtaposition to each other. The fact of their existence in harmony is what is meaningful. But "meaningful" to what? To the *mind*.

According to the hypothesis offered here, the term "meaning" is another way to talk about the ability of the mind to juxtapose its ideas. Sometimes these juxtapositionings are harmonious (meaningful) and sometimes they are not.

REFERENCES

1. BLANSHARD, B. Reason and Analysis. LaSalle, Ill.: Open Court, 1962.
2. LAWSON, C. A. Language, Thought, and the Human Mind. East Lansing, Mich.: Michigan State Univ. Press, 1958.
3. OATES, W. J. The Communication of Ideas (ed. by L. Bryson). New York: Harper, 1948.
4. STEWART, D. K. Communication, ideas, and meaning. *Psychol. Rep.*, 1965, **16**, 885-892.
5. ———. Signs, symbols, and meaning. *J. Communication,* 1966, **16**, 4-9.
6. ———. Communication and ideas. *J. Amer. Sci. Affiliation* (in press).
7. WERKMEISTER, W. H. The meaning of "Meaning" re-examined. *Philos. Rev.*, 1938, **47**, 245-266.
8. WHORF, B. L. Language, thought, and reality. In J. B. Carroll (Ed.), *Selected Writings*. Cambridge, Mass.; M. I. T. Press, 1964.

Research Department
Campbell-Ewald Company
General Motors Building
Detroit, Michigan 48202

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 101-105.

HEIGHT, WEIGHT, AND PONDERAL INDEX OF COLLEGE MALE SMOKERS AND NONSMOKERS*

Department of Psychology, Boston College

DANIEL J. BAER

A. INTRODUCTION

Although there is some evidence that smokers differ from nonsmokers in body shape, so far the results have not been conclusive (10). Seltzer (7) reports that for a 13-year longitudinal study of Harvard students, smokers were consistently greater than nonsmokers in height and weight, as well as other body dimensions. Thomas (11), using a sample of 1000 medical college students, found that smokers, when compared to nonsmokers, tended to have unusually heavy body builds. However, Damon (2), on the basis of a sample of 167 adult male factory workers of Neapolitan parentage, concluded that lean men tend to smoke more than fat, but not muscular men.

So far, most of these studies have either employed univariate statistical analysis of body dimensions (4, 7, 11) or else have reduced these multiple measurements into some taxonomy (2, 5, 7). However, since even the same taxonomy is frequently not comparable from laboratory to laboratory (3), and because body measurements tend to be correlated, multivariate statistical procedures may be more appropriate methods of establishing group differences in body shape. The purpose of the present study was to compare Heavy Smokers, Light-Moderate Smokers, Former Smokers, and Nonsmokers in their body height, weight, and ponderal index by means of both univariate and multivariate statistical procedures.

B. METHOD

1. *Subjects*

The *S*s were 405 white, American-born, Catholic college students. Heavy Smokers ($N = 143$) smoked at least one pack of cigarettes a day, or its equivalent in pipes and cigars. Light-Moderate Smokers ($N = 106$) smoked less than a pack a day, while Former Smokers ($N = 32$) were once habitual smokers, but had stopped smoking for at least one month prior to the interview.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on July 19, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

The Nonsmokers ($N = 124$) were those individuals who had never smoked, except for a possible rare trial experience.

2. *Procedure*

As part of a smoking behavior survey (1), *S*s reported their height, while barefoot, in inches; their weight, without clothes, in pounds; and their age, rounded to the closest year. Ponderal index (9) was subsequently determined by dividing the height by the cube root of the weight.

C. RESULTS AND DISCUSSION

Table 1 summarizes the mean and standard deviation of the age, weight, height, and ponderal index of the four groups. The distribution for each of these variables approximated a unimodal leptokurtic normal distribution with comparable means and standard deviations (*SD*s) for the four groups. Heavy Smokers were one-half inch taller, and Former Smokers were one-half inch shorter, than Light-Moderate Smokers and Nonsmokers.

TABLE 1
MEAN AND STANDARD DEVIATION OF AGE, WEIGHT, HEIGHT, AND PONDERAL INDEX OF FOUR GROUPS DIFFERING IN SMOKING BEHAVIOR

Group	*N*	Age (years) mean	Age (years) *SD*	Weight (lbs.) mean	Weight (lbs.) *SD*	Height (inches) mean	Height (inches) *SD*	Ponderal index mean	Ponderal index *SD*
Heavy Smokers	143	20.5	1.17	169	22.3	70.9	2.34	13.1	0.14
Light-Moderate Smokers	106	20.4	1.00	166	23.5	70.2	2.62	13.0	0.16
Former Smokers	32	20.7	1.14	164	19.1	69.7	2.48	12.9	0.15
Non-smokers	124	20.5	1.00	166	23.4	70.2	2.44	13.0	0.17

For each variable, the *t*-tests of the differences between means of the groups were insignificant except for body height. In Table 2, it may be seen that Heavy Smokers were significantly ($p < .05$) taller than Light-Moderate Smokers, Former Smokers, and Nonsmokers. Since body weight was similar for these four groups, it would seem that Heavy Smokers were more ectomorphic in body shape than other individuals. The higher mean ponderal index of Heavy Smokers, which according to Sheldon (9) is an excellent measure of ectomorphy for young adults, is consistent with such a suggestion. Thus, this outcome would support Damon's conclusion (2) that leaner men tend to be heavier smokers. Perhaps after 10 to 15 years, when weight differences become more pronounced while height remains relatively constant, ponderal index may more clearly distinguish these four groups.

Although insignificant, it is interesting that Former Smokers tended to be shorter and heavier than the other groups. Perhaps a short fat man may more readily relinquish the smoking habit than a tall thin man. Such a finding, if substantiated, could be useful information to smoking clinics in the type of training methods selected for their clients.

Because height, weight, ponderal index, and age are interrelated and yet each measure represents a somewhat different body characteristic, it would seem that a generalized T test (6) might reveal differences not observed by the univariate *t*-tests. Results for the stepwise discriminate function analysis of the Heavy Smokers compared to each of the other three smoking groups are reported in Table 3. When the vector of means for the four variables were compared, Heavy Smokers did not differ from Light-Moderate Smokers, Former Smokers, and Nonsmokers. With each step, the one variable that contributed the least to the discrimination was removed from the remaining variables. For step 4, for the Light-Moderate Smokers and Nonsmokers, only body height significantly differentiated the groups ($p < .05$ one-tailed hypothesis). This result is necessarily the same as that observed with the univariate *t*-test. Since $F = t^2$ for one *versus* n_2 degrees of freedom, the F's in Table 3 are the same as the t's in Table 2. However, when the Heavy Smokers were compared to the Former Smokers, height was the first variable eliminated from the four measures and the vector based upon ponderal index and weight, although insignificant, offered the greatest discrimination between the groups. Probably the great difference in sample sizes, as well as lack of homogeneity of the variance-covariance matrix, contributed to this outcome, so that the univariate *t*-test detects differences not apparent by the T^2 test. Thus, in future studies, it would seem that both types of procedures would be useful.

TABLE 2

t TEST OF THE DIFFERENCE BETWEEN MEAN HEIGHT OF HEAVY SMOKERS COMPARED TO THE MEAN HEIGHT OF THREE OTHER GROUPS

Group	Light-Moderate Smokers	*t* Former Smokers	Nonsmokers
Heavy Smokers	1.89*	2.52**	1.89*

* $p < .05$ (one-tailed hypothesis).
** $p < .01$ (one-tailed hypothesis).

In order to determine if height differences exist between Heavy Smokers and other groups when the effects of age and body weight were removed, an analysis of covariance was performed. Table 4 summarizes the adjusted and unadjusted *F*'s, based upon these analyses. The unadjusted *F*'s represent the outcomes of the analysis of variance of height differences between the Heavy Smokers and each of the other three groups without regard to body weight or age, while the adjusted *F*'s were based upon the analysis of covariance of height differences when the effects of these two variables were statistically removed. It may be seen that the outcome for covariance analysis was the same as analysis of variance for the Nonsmokers and Former Smokers, but not for the Light-Moderate Smokers. These results indicate that weight and age are not important factors accounting for height differences of these groups, except when Heavy Smokers were compared to Light-Moderate Smokers. From a separate analysis of covariance of height differences, with first age differences alone and then weight differences alone removed, it was found that age primarily contributed to the insignificant result. Presumably, the

TABLE 3
STEPWISE DISCRIMINATE FUNCTION ANALYSIS OF HEAVY SMOKERS COMPARED TO EACH OF THREE OTHER SMOKING GROUPS BASED UPON HEIGHT, WEIGHT, AGE, AND PONDERAL INDEX

Heavy Smokers compared to	Step	*df*	*F*	Variable eliminated
Light-Moderate Smokers	1	4,244	1.34	none
	2	3,245	1.50	age
	3	2,246	1.79	age, weight
	4	1,247	3.59*	age, weight, ponderal index
Nonsmokers	1	4,262	1.49	none
	2	3,263	1.92	age
	3	2,264	2.43	age, ponderal index
	4	1,265	4.58*	weight, age, ponderal index
Former Smokers	1	4,170	1.92	none
	2	3,171	2.49	height
	3	2,172	3.42	height, age
	4	1,173	7.41	height, age, weight

* $p < .05$ (one-tailed hypothesis).

TABLE 4
ANALYSIS OF COVARIANCE OF BODY HEIGHT OF HEAVY SMOKERS WHEN COMPARED TO EACH OF THREE OTHER GROUPS WITH THE EFFECTS OF AGE AND BODY WEIGHT REMOVED

	Unadjusted height		Adjusted height	
Heavy Smokers compared to	*df*	*F*	*df*	*F*
Light-Moderate Smokers	1,247	3.59*	1,245	2.37
Nonsmokers	1,265	4.58*	1,263	4.12*
Former Smokers	1,173	6.34*	1,171	4.70*

* $p < .05$ (one-tailed hypothesis).

Light-Moderate Smokers, with increasing age, may not only increase in height, but also more likely may become Heavy Smokers.

Although, in general, the same conclusions were reached regardless of statistical procedure, the present analysis does demonstrate that each method can contribute unique advantages. Thus, the important suggestion by Domey *et al.* (3), that multivariate techniques be applied when studying body shape differences, should not encourage investigators to overlook the merits of the well established and sometimes more easily understood univariate inferential statistics.

D. SUMMARY

A sample of 143 Heavy Smokers, 106 Light-Moderate Smokers, 32 Former Smokers, and 124 Nonsmokers was compared in their body height, weight, and ponderal index. Although no differences were found among the groups in their weight and ponderal index, Heavy Smokers were significantly taller (.05 level) than the other three groups. Former Smokers were one-half inch shorter, while Heavy Smokers were one-half inch taller than Light-Moderate Smokers and Nonsmokers. Univariate as well as multivariate statistical procedures revealed similar outcomes.

REFERENCES

1. BAER, D. J. Smoking attitude, behavior, and beliefs of college males. *J. Soc. Psychol.*, 1966, **68**, 65-78.
2. DAMON, A. Constitution and smoking. *Science*, 1961, **134**, 339-341.
3. DOMEY, R. G., DUCKWORTH, J. E., & MORANDI, A. J. Taxonomies and correlates of physique. *Psychol. Bull.*, 1964, **62**, 411-426.
4. HEATH, D. W. Differences between smokers and nonsmokers. *AMA Arch. Intern. Med.*, 1958, **101**, 377-388.
5. PARNELL, R. W. Smoking and cancer. *Lancet*, 1951, **1**, 963.
6. RAO, C. R. Advanced Statistical Methods in Biometric Research. New York: Wiley, 1952.
7. SELTZER, C. C. Masculinity and smoking. *Science*, 1959, **130**, 1706-1707.
8. ———. Morphologic constitution and smoking. *J. Amer. Med. Assoc.*, 1963, **183**, 639-645.
9. SHELDON, W. H. The Varieties of Human Physique. New York: Harper, 1940.
10. THE SURGEON GENERAL'S ADVISORY COMMITTEE ON SMOKING AND HEALTH. Smoking and Health. Washington, D.C.: U.S. Government Printing Office, 1964.
11. THOMAS, C. B. Characteristics of smokers compared with nonsmokers in a population of healthy young adults, including observations on family history, blood pressure, heart rate, body weight, cholesterol, and certain psychological traits. *Ann. Intern. Med.*, 1960, **53**, 697-718.

Department of Psychology
Boston College
Chestnut Hill, Massachusetts

Light-Moderate Smokers, [illegible] not only [illegible] in height, but also more likely [illegible] Smokers.

Although in general [illegible] constitutional [illegible] of [illegible] procedures, the present [illegible] unique advantages. Thus, the [illegible] by [illegible] (12), [illegible] techniques be applied [illegible] should not [illegible] the [illegible] well established and sometimes [illegible] experimental [illegible].

[illegible] Summary

A sample of [illegible] Heavy Smokers, [illegible] Light-Moderate Smokers, [illegible] Ex-Smokers, and [illegible] Nonsmokers was compared in their body height, weight, and ponderal index. Although no differences were found among the groups in their weight and ponderal index, Heavy Smokers were significantly [illegible] than the other three groups. [illegible] Smokers were one half inch shorter, while [illegible] Smokers were one half inch taller than Light-Moderate Smokers and Nonsmokers. [illegible] as well as [illegible] results should be [illegible].

References

1. [illegible], D. [illegible] Smoking attitude, behavior, and [illegible] of college students. J. [illegible], 1966, [illegible].
2. [illegible], A. Constitution and smoking. Science, [illegible].
3. [illegible], [illegible] & [illegible] [illegible].
4. [illegible], D. W. Differences between smokers and nonsmokers. [illegible].
5. [illegible], R. W. [illegible]. Lancet, [illegible].
6. [illegible], G. [illegible] Statistical Methods [illegible]. New York: Wiley, [illegible].
7. [illegible], C. [illegible] and smoking. Science, [illegible].
8. [illegible] Morphologic constitution and smoking. J. Amer. Med. Assn., [illegible].
9. [illegible], W. H. The Varieties of Human Physique. New York: Harper, [illegible].
10. [illegible] Report of the Advisory Committee to the Surgeon General of the Public Health Service. Washington, D.C.: U.S. [illegible].
11. [illegible], G. [illegible] smoking [illegible] population [illegible] body weight [illegible], [illegible].

Department of Psychology
[illegible] College
[illegible], Massachusetts [illegible]

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 107-112.

ORIGINALITY IN PROBLEM SOLVING AS AFFECTED BY SINGLE- *VERSUS* MULTIPLE-SOLUTION TRAINING PROBLEMS*[1]

Cognitive Operations Laboratory, Bethany College

WILBERT S. RAY

A. INTRODUCTION

Originality in thinking may be defined as the number of answers produced in any problem or as responses of a particular kind, as when Maltzman *et al.* (6) count unique or nearly unique items or Mednick (2) gives credit only for answers that meet the specifications of his Remote Associates Test. The present experiment counted all responses to the test problem. The work was with problem-solving situations rather than with the verbal associations used by Maltzman and Mednick.

The hypothesis tested was that training with problems requiring multiple solutions would facilitate multiple solutions to a test problem more than would training with problems requiring only single solutions.

B. EXPERIMENT I

1. *Method*

a. Tasks. Five problems were used, the two first described being single-solution problems:

(1). *Triangles. S* is given a triangle, as shown in Figure 1, and asked

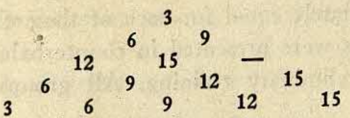

FIGURE 1
TRIANGLE TASK USED IN EXPERIMENT I

what number belongs in the empty space. Ten such problems were used, the first and the last being very easy. (Norms for these triangle tasks had been established previously.)

---

* Received in the Editorial Office, Provincetown, Massachusetts, on July 25, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was partially supported by Public Health Service Grant MH 08823-02.

(2). *Making the last draw* (8). Ten paper clips are laid on a table; *S* is given the first draw (of one or two objects), and must learn to get the last one.

(3). *Miniature Maier two-string task* (4). A box, 21 inches long, 9 inches high, and 9 inches deep, stands on a table, an open side toward *S*. Two strings hang from the ceiling to within an inch of the floor. A doll's chair, 2½ inches high, stands in a corner. On the table beside the box lie, full-sized, a meter stick, a ball of cord, a box of safety matches, a pair of pliers, two large paper clips, a candle, a right-angle clamp, and a book. *S* is asked to imagine that the room is full-sized, and that the strings are far enough apart that he can not reach both of them. He is requested to tell *E* what procedure he could use to tie the strings together, using any of the available objects; and is asked for as many solutions as possible.

(4). *Brick uses* (1). *S* is requested to give *E* as many uses as possible for a brick.

(5). *The cylinder in the can* (10). A metal cylinder (the "can"), 3 3/4 inches high by 1 3/8 inches inside diameter, is welded to a base. A wooden cylinder, 2 inches high and with an outside diameter of 1 1/4 inches, is handed to *S*, then put into the can. *S* is requested to think of ways to get it out. This was the test problem in all treatment conditions.

b. Procedure. Four training conditions were used. The Single Solution group (SS) worked on the two single-solution problems, the Multiple Solution group (MS) on the two multiple-solution problems. The Listening group (L) listened (5) to solutions of the multiple-solution problems rather than producing their own solutions—nine for the two-string and 18 for the brick-uses tasks. These numbers were chosen in an attempt to make the elapsed time approximately equal for each of these three conditions. In each case the two problems were presented in counterbalanced sequences over *S*s. Group C had no preliminary training. All groups were tested with the cylinder in the can.

A total of 48 *S*s were used, 12 in each condition. All were students in general psychology who participated as part of their class work.

Four students from a course in experimental psychology acted as *E*s, each in a separate room, each running three *S*s in each treatment. Treatments were counterbalanced over *E*s, and *S*s were assigned to *E*s at random with a restriction as to equal numbers.

2. *Results*

All solutions to the test problem were counted, whether "correct" or not. Mean numbers of solutions by conditions and by *E*s are shown in Table 1, along with the analysis of variance. Treatment groups differ from each other at the .05 level. Comparisons of the differences between treatments in terms of t appear in Table 2, where the SS group differs from the other groups at the .02 level.

The prediction was that the MS group would produce more solutions than would either the SS or the C group. Neither part of the prediction was supported. Indeed, since the SS group scored reliably lower than any of the other groups, the only conclusion possible is that the single-solution training seems to have inhibited originality as here defined.

TABLE 1

ANALYSIS OF VARIANCE OF DATA FROM EXPERIMENT I

Source	*df*	*MS*	*F*	*p*
Treatments	3	49.57	3.00	$<.05$
Experimenters	3	51.07	3.09	$<.05$
Interaction	9	10.61		
Error	32	16.54		

Note: The mean numbers of solutions by treatment groups were as follows: Single Solution group, 9.25; Multiple Solution group, 12.00; Listening group, 13.92; Control group, 13.08; and by Experimenters, 12.25, 11.17, 10.00, and 14.83.

TABLE 2

DIFFERENCES AMONG TREATMENT MEANS

	Group		
Group	Multiple Solution	Listening	Control
Single Solution	2.75*	4.67*	3.83*
Multiple Solution		1.92	1.08
Listening			.84

* With 22 *df*, t is significant at .02 or better.

There is also a reliable difference among *E*s, but no interaction of *E*s with treatments.

3. *Discussion of Experiment I*

The different results secured by the *E*s, but with no interaction with treatments, present a problem to which there is no easy answer. It is reasonable to assume that some *E* or *E*s produced more reliable results than did others, but there is no way to identify them. It is obviously no solution to use only one *E*, even if he were chosen at random from the four. More experience and training for the *E* may be suggested, but there is a growing suspicion today (3, 7, 9) that no *E* can be depended on not to influence the results. Another solution may lie in the type of experiment reported by Videbeck and Bates (11), who had a computer give the instructions to *S*, present the tasks and reinforcement, and record and count the responses. The question of the reliability of results so secured has not yet been raised.

It might be predicted that listening to multiple solutions would show less facilitation of the originality than would producing multiple solutions. No such decrement was found. Maltzman, Bogartz, and Breger (5) found that a request for originality actually produced more originality than did training. The question arises as to whether it is possible that one source of originality lies in the *S*'s expectancy as to the number of solutions he will find. Expectancy appears to be a sort of set.

C. EXPERIMENT II

The expectancy hypothesis led to another experiment in which the independent variable consisted of statements to *S*s as to the number of solutions that were usually produced.

All *S*s were given a sheet of paper that carried the sentence, "A safety pin is ordinarily used for fastening things together. Name other possible uses for this safety pin" (6). The *S*s, students from a course in general psychology who participated as part of their laboratory work, were randomly divided into two groups. The *E* (only one *E*, the present writer) said to one group, "The average number of answers is six, and I hope you will do better than that." The other group was told the average number was 15. The "six" group produced a mean of 9.03 answers, the "15" group gave 11.43, a difference significant at the .05 level with 57 *df*.

The experiment supported the expectancy hypothesis.

D. DISCUSSION

It would seem that there may be two ways to facilitate originality defined as number of solutions: practice at producing many solutions, and the building of an expectancy of many solutions. Instructions to *S* to be original [see Maltzman *et al.* (5)] should probably be classified as a method of producing an expectancy. Practice with multiple-solution tasks will itself produce an expectancy of multiple solutions.

In Experiment I all groups received with the test problem the words, "Tell me as many ways to get it out as you can think of. Don't omit any of them because they are too obvious." This statement should produce an expectancy of many solutions in every group, which may explain why the control group produced as many solutions as did any other group. However, the single-solution training group produced fewer responses than did any other group; which suggests that the low expectancy that resulted from training with single solutions was stronger than the high expectancy resulting from the test-problem instructions.

It may also be concluded that the facilitation assumed to result from the multiple-solution training, or from listening to multiple solutions, did not raise the amount of originality above that produced by the test-problem instructions. In fact, there is here no basis for a conclusion that the practice with multiple solutions had any effect at all. The hypothesis, however, as to the effects of practice, is so reasonable and attractive that it should not be rejected on the basis of any one experiment.

It will be noted that it is easier to design an experiment to produce expectancies without practice than an experiment with practice but without expectancies. It should be easy, however, to use a test task that is not immediately preceded by high-expectancy instructions.

E. Summary

The hypothesis tested was that practice with problems having many solutions would produce more originality than would practice with problems requiring only one solution. In addition to the two practice groups, a third group listened to multiple solutions of practice problems, and a control group had no preliminary training. Originality was defined as the production of many solutions to a test problem. The only statistically supported conclusion is that the single-solution training problems depressed the amount of originality.

Examination of the results suggested that *S*'s expectancy of many solutions may produce many solutions. A second experiment supported this hypothesis.

References

1. Guilford, J. P., Wilson, R. C., Christensen, P. R., & Lewis, D. J. A factor-analytic study of creative thinking: I. Hypotheses and description of tests. Reports from the Psychological Laboratory, No. 4, University of Southern California, Los Angeles, California, 1951.
2. Houston, J. P., & Mednick, S. A. Creativity and the need for novelty. *J. Abn. & Soc. Psychol.*, 1963, **66**, 137-141.
3. Kintz, B. L., Delprato, D. J., Mettee, D. R., Persons, C. E., & Schappe, R. H. The experimenter effect. *Psychol. Bull.*, 1965, **63**, 223-232.
4. Maier, N. R. F. Reasoning in humans: I. On direction. *J. Comp. Psychol.*, 1930, **10**, 115-143.
5. Maltzman, I., Bogartz, W., & Breger, L. A procedure for increasing word association originality and its transfer effects. *J. Exper. Psychol.*, 1958, **56**, 392-398.
6. Maltzman, I., Simon, S., Raskin, D., & Licht, L. Experimental studies in the training of originality. *Psychol. Monog.*, 1960, **74**, No. 493.
7. Orne, M. T. On the social psychology of the psychological experiment. *Amer. Psychol.*, 1962, **17**, 776-783.
8. Peterson, J. C. The higher mental processes in learning. *Psychol. Monog.*, 1920, **28**, No. 129.
9. Pfungst, O. Clever Hans (ed. by R. Rosenthal). New York: Holt, Rinehart and Winston, 1965. Pp. ix-xxxiv.

10. **Saugstad, P.** Incidental memory and problem solving. *Psychol. Rev.,* 1952, **59,** 221-226.
11. **Videbeck, R.,** & **Bates, H.** Verbal conditioning by a simulated experimenter. *Psychol. Rec.,* 1966, **16,** 145-152.

Department of Psychology
Cognitive Operations Laboratory
Bethany College
Bethany, West Virginia 26032

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 113-118.

SOCIAL DESIRABILITY, EXTRAVERSION, AND NEUROTICISM: A LEARNING ANALYSIS*

University of London

FRANK H. FARLEY

A. INTRODUCTION

In an analysis of Social Desirability (SD) in the Maudsley Personality Inventory (MPI) (6), Martin and Stanley (14) found that the Neuroticism (N) Scale had a significant negative correlation of —.61 with their own SD Scale, but that the Extraversion (E) Scale did not contain SD variance. The Lie (L) Scale had a significant positive correlation of .39 with their SD Scale.

The MPI has since been thoroughly restandardised and refactored with alterations in item content (7) and has been named the Eysenck Personality Inventory (EPI). The present paper reports an analysis of SD responding in Extraversion and Neuroticism as measured by the EPI.

In the Martin and Stanley study, SD was assessed by an SD Scale constructed by the authors. However, the most widely used scale, that of Edwards (4), was not employed. In light of the large body of literature on the Edwards Scale (ESD), it would seem worthwhile employing this instrument when investigating relationships between SD and personality inventories.

The ESD is based on the conceptualization that a socially desirable response is a True response to any item with a high social desirability scale value (as assigned by judges) and a False response to any item with a low social desirability scale value. The ESD has been constructed in terms of this definition of SD responding. It consists of 39 items drawn from the MMPI with nine true-keyed items of high social desirability scale value, and 30 false-keyed items of low social desirability scale value. Edwards' approach to SD research is essentially an item-bound one; it concerns itself with the scale-value of items along a judged Social Desirability dimension.

Another approach to SD research is represented by Crowne and Marlowe (3), who have contrasted the purely psychometric analysis of SD with the investigation of a subject's *motivation* to respond in the socially desirable direction, and the extent to which this behavior generalizes to nontest situations. They have interpreted SD responding as a reflection of a subject's need

---

* Received in the Editorial Office, Provincetown, Massachusetts, on July 25, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

for social approval. In an extensive series of studies, they have demonstrated that SD responding in testing situations can be predictive of a subject's conforming, approval-seeking behavior in a wide range of nontest situations. They have criticized the ESD as a measure of Social Desirability on the grounds that it is contaminated by psychopathological item content. That Edwards has demonstrated consistently high correlations between the ESD and MMPI Scales may be due to the effects of SD upon endorsement of items in the latter, or simply be due to the psychopathology in both scales. It is not possible to sort out SD response bias from genuine endorsement of symptomatology. Crowne and Marlowe (2) have developed an SD Scale (MC-SD) that is supposedly free of pathological item content and is based on a generalized motivational interpretation of SD responding. They consider that the MC-SD reflects more than SD response bias; it is a measure of the personality trait need-for-approval. Martin and Stanley (14) have emphasized that SD seems to be a common factor in effective pathological scales, and that it may be a requisite of such scales.

Heilbrun (12) has reviewed the whole issue of SD and psychopathology, and has proposed that psychological health and Social Desirability are "in large measure" one and the same thing. ". . . it becomes important to distinguish between the tendency to behave in socially desirable ways, which must be mirrored in test performance to the extent veridical self-appraisal is achieved, and the tendency to respond to test items without regard to fact but only with regard to the social appearance which the person wishes to create by his test performance" (12, p. 385). Block (1) distinguished between a facade or superficial brand of socially desirable behavior, and socially desirable behavior that is a correlate of adjustment. Heilbrun identified Block's first type as a "fake good" set, and the second as a basic product of social learning.

In an earlier study, Heilbrun (11) demonstrated that the more socially undesirable were the keyed responses on any MMPI Scale, the better that scale was at differentiating psychologically disturbed *S*s from normal ones. "Since the pathological groups invariably received the higher scale scores, it follows that they were endorsing the keyed, socially undesirable and psychopathological behavior more frequently. Therefore, the more socially undesirable the behavior endorsed by an individual, the more likely he will evidence a manifest psychological disturbance" (12, p. 384).

On the basis of the above discussion, one might consider SD responding and the Extraversion and Neuroticism Scales as follows. Any correlations between SD Scales and N Scales could be due to shared psychopathological item content, or essential SD variance reflecting the respondent's adjustment to social and cultural norms, and, therefore, his learning history. From Crowne and Marlowes' (2) analysis, it would be expected that ESD and N correlations would largely be a product of similarity of pathological-type content. But correlations between their scale (MC-SD) and the N Scale would be interpreted as reflecting the presence or absence of socially conditioned, approval-motivated traits (indicative of positive adjustment) along the neuroticism dimension. The latter relationship would seem to represent Heilbrun's equation of psychological health and Social Desirability. To the extent that an ESD and N correlation might reflect the operation of at least both reasons outlined above, the ESD would be expected to have a higher negative correlation with the N Scale than would the MC-SD. The MC-SD would, in effect, be controlling for that confounding due to similarity of pathological item content.

As there is no psychopathological item content in the EPI-E Scale, and Martin and Stanley found no relationship between their SD Scale and the MPI-E, it was expected that no significant correlation would be found between the EPI-E and ESD. Where the MC-SD is concerned, as it seems to refer less to the SD characteristics of questionnaire items and more to a socialized, conforming, approval-motivated characteristic of personality, it is conceivable on the basis of the personality-behavior theory of Eysenck (5) that, to the extent that the extravert is less socially conditioned than the introvert, he will respond less in the approval-motivated direction. That is, the EPI-E would correlate negatively with the MC-SD. The stronger social conditioning of the introvert would be reflected in his endorsement of socially desirable items as contained in the MC-SD. Using the same learning-conditioning analysis, a negative correlation would be expected between the N Scale and MC-SD. This would represent the neurotic's endorsement of socially undesirable items as simply one part of his maladaptive (asocial) behavior pattern; he is reflecting a history of maladjustive learning.

B. METHOD

One hundred male subjects were used. Sixty-eight of these were apprentices at a large motor works near London, and the remaining 32 were paid volunteers taking part in a number of experiments at the Institute of Psychiatry, University of London. These latter subjects were largely drawn from Civil Service occupations in the greater London area. In conjunction with other studies, the subjects were administered Form A of the EPI, the Edwards' 39-item SD Scale, and the Marlowe-Crowne 33-item SD Scale. The EPI was scored for E, N, and L.

C. Results and Discussion

Product-moment correlations were computed between the SD Scales and the EPI Scales. These correlations were as follows: ESD—Extraversion, $r = .133$; ESD—Neuroticism, $r = —.603$ ($p < .01$); ESD—Lie, $r = .069$; MC-SD—Extraversion, $r = —.218$ ($p < .05$); MC-SD—Neuroticism, $r = —.299$ ($p < .01$); and MC-SD—Lie, $r = .554$ ($p < .01$).

It can be seen that, with 98 *df*, the ESD correlates significantly only with the N Scale, but the MC-SD correlates significantly with all EPI Scales.

It should be noted here that of the 33 items comprising the MC-SD, 18 are keyed true and 15 false, so that an interpretation of MC-SD correlations in terms of an acquiescence response set is highly improbable. This may not be asserted with confidence where the ESD is concerned, as this scale contains nine items keyed true and 30 items keyed false. Hand and Brazzell (10) have reported data suggesting that the ESD does, and the MC-SD does not, harbor acquiescence variance.

Where the correlations between extraversion and the SD Scales are concerned, the lack of a significant correlation between the E Scale and the ESD clearly suggests that the E Scale is free of SD variance in the symptomological sense. This suggestion is reinforced by the high negative correlation between the N Scale and the ESD. The very considerable drop in the magnitude of correlation between the N Scale and the MC-SD when compared to that between the N Scale and the ESD, coupled with the lack of correlation between E and ESD, strongly suggests that the ESD is contaminated to a considerable extent with psychopathological item content, and that correlations between the ESD and pathological scales cannot be interpreted as necessarily reflecting SD variance in the latter.

The negative correlation of the E Scale with the MC-SD is in accord with a social conditioning analysis of Extraversion, and suggests that points of contact may be made between Eysenck's personality theory and recent work on approval motivation (3). As the correlation of E and MC-SD is low but of particular interest where the theoretical analysis of SD responding is concerned, replication of this relationship seemed desirable. Accordingly a further 58 male apprentices were administered the E Scale and MC-SD. In this replication the correlation was —.38 ($p < .01$). It is apparent that the correlation has held up, and indeed is much higher.

The correlations between the N Scale and the ESD and MC-SD are in line with the considerations advanced earlier; the correlation of —.603 between N and ESD most likely reflects similarity of pathological item content; the correlation of —.299 between N and MC-SD most likely mirrors essential "adjustive" SD variance (the learning analysis), and may possibly also reflect a slight degree of pathological content in the MC-SD (13).

The .069 correlation between the EPI-L and ESD is somewhat lower than the usual insignificant, positive correlations between the ESD and Lie Scales as reported in the literature (3), and strongly suggests that the ESD is not unambiguously measuring an individual's tendency to describe himself in a socially desirable light. That the EPI-L is largely an SD measure has been demonstrated by Eysenck and Eysenck (8) employing a preliminary form of that scale. The correlation of .554 between the EPI-L and MC-SD is virtually identical with the correlation between the MMPI-L Scale and MC-SD of .54 reported by Crowne and Marlowe (3) for a sample of 37 male and female American college students, and with the .57 correlation between the MC-SD and the Lie Scale of the Maudsley Personality Inventory found by Farley (9) for a sample of 70 male Canadian university students, and the unpublished correlation of .49 between the MC-SD and the MPI-L Scale found by the writer for 30 male and female Canadian introductory psychology students.

D. SUMMARY

On the hypotheses that introverts socially condition more strongly than extraverts and should therefore emit more socially desirable (SD) responses than extraverts, and that *S*s high in Neuroticism (N) should endorse fewer SD responses than low N *S*s as a reflection of the neurotic's maladaptive learning history, measures of Extraversion (E) and N (Eysenck Personality Inventory) were correlated on 100 male English *S*s with two measures of SD responding (Edwards SD Scale and Marlowe-Crowne SD Scale). In line with expectation, both E and N were significantly negatively associated with SD responding. There was, however, no significant relationship between E and the Edwards scale. The magnitude of the relationships between SD and each of the personality dimensions was highly dependent on the SD measures employed, and was consistent with theoretical analyses of these measures.

REFERENCES

1. BLOCK, J. Some differences between the concepts of social desirability and adjustment. *J. Consult. Psychol.*, 1962, **26**, 527-530.
2. CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, **24**, 349-354.
3. ———. The Approval Motive: Studies in Evaluative Dependence. New York: Wiley & Sons, 1964.
4. EDWARDS, A. L. The Social Desirability Variable in Personality Assessment and Research. New York: Dryden, 1957.

5. EYSENCK, H. J. The Dynamics of Anxiety and Hysteria. London: Routledge & Kegan Paul, 1957.
6. ———. Manual of the Maudsley Personality Inventory. London: Univ. London Press, 1959.
7. EYSENCK, H. J., & EYSENCK, S. B. G. Manual of the Eysenck Personality Inventory. London: Univ. London Press, 1964.
8. EYSENCK, S. B. G., & EYSENCK, H. J. An experimental investigation of "desirability" response set in a personality questionnaire. *Life Sci.*, 1963, **5**, 343-355.
9. FARLEY, F. H. Global self-ratings, the independence of questionnaire drive and anxiety, and social desirability variance. Unpublished manuscript, Institute of Psychiatry Library, University of London, London, 1965.
10. HAND, J., & BRAZZELL, C. O. Contamination in measures of acquiescence and social desirability. *Psychol. Rep.*, 1965, **16**, 759-760.
11. HEILBRUN, A. B. Revision of the MMPI K correction procedure for improved detection of maladjustment in a normal population. *J. Consult. Psychol.*, 1963, **27**, 161-165.
12. ———. Social-learning theory, social desirability and the MMPI. *Psychol. Bull.*, 1964, **61**, 377-387.
13. KATKIN, E. S. The Marlowe-Crowne social desirability scale: Independent of psychopathology? *Psychol. Rep.*, 1964, **15**, 703-706.
14. MARTIN, J., & STANLEY, G. Social desirability and the Maudsley Personality Inventory. *Acta Psychol.*, 1963, **21**, 260-264.

Department of Educational Psychology
The University of Wisconsin
936 West Johnson Street
Madison, Wisconsin 53706

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 119-126.

THE PHONETIC CONTENT OF THE CRIES OF NORMAL INFANTS AND THOSE WITH BRAIN DAMAGE*[1]

The Long Island Jewish Hospital

VINCENT R. FISICHELLI, MALCOLM COXE, LAURA ROSENFELD, AUDREY HABER, JOAN DAVIS, AND SAMUEL KARELITZ[2]

A. INTRODUCTION

Previous studies of the cries of normal infants and infants with diffuse brain damage have revealed significant differences between the two groups in terms of threshold (6) and latency (2). In addition, the cries of normal infants are more rhythmic and they show earlier production of inflectional changes than do the cries of abnormal infants (7).

The present study represents a further effort to differentiate between the cries of normal and abnormal infants: in this instance, in terms of the phonetic structure of the cry. Similar surveys have already been made by Irwin (3) for speech sounds, by Fisichelli (1) for the babbling sounds of institutionalized infants, and by Irwin and his associates for the cries of normal infants (4, 5).

B. SUBJECTS

The samples of crying that were subjected to phonetic analysis were taken from the Long Island Jewish Hospital library of tape recordings of normal and brain-damaged infant vocalizations. Tapes of 10 abnormal infants randomly selected at each of three age levels—the first week of life (one to eight days), approximately six months (172 to 229 days), and close to one year (342 to 393 days)—were matched with tapes of normal subjects according to age and sex.

Two of the subjects (Numbers 18 and 59)[3] in the abnormal group were found to have *IQ*s of 87 and 86, respectively, three years after this study was completed. This finding introduced the possibility of excluding them from the abnormal group. One of these infants was recorded on the eighth day of life. He had sustained a subarachnoid hemorrhage at birth and because of twitching was placed on phenobarbitol. Apparently the recovery was considerable, since his motor development was normal at age three years and his *IQ* was 87. Thus, this infant fits into both categories: brain-disturbed in early infancy, and almost normal later. The second child was flabby and poorly coordinated in early infancy. He was then considered to be severely brain-damaged, yet he improved with time and had an *IQ* of 86 at age three years. Here, again, there may have been birth injury with subsequent recovery.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on July 25, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study is supported by Grant HD00332 of the National Institutes of Health, Division of Research Grants.

[2] The authors are grateful to Miss Amalie Ansorge for her generous assistance.

[3] The data which are summarized here are available at ADI for each of the 60 subjects studied. Material supplementary to this article has been deposited as Document number 9024 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. A copy may be secured by citing the Document number and by remitting $1.25 for photoprints, or $1.25 for 35-mm micro-

All of the statistics reported in this study were computed with and without the scores of these two infants. With the exception of mean consonant frequency discussed below, the exclusion or inclusion of these two cases made no difference in the results obtained; therefore, all statistics are reported for the entire group.

C. Procedure

The primary object of the stimulation procedure was to elicit crying in the infant, who was not found crying at the time of recording, by flicking the sole of the foot with the experimenter's index finger.

All test sessions were recorded on magnetic tape with an Ampex 601 (Electrovoice 666 microphone), and appropriate data, including time of day, last feeding time, etc., were preserved.

The tapes subjected to phonetic analysis were re-recordings that were carefully matched for intensity levels with the original sample. These tapes were stripped of all identifying data or information that might convey identity to the phonetician.[4] The samples were arranged on continuous reels in such a manner as to preclude any practice effect for any given age group or for normal or abnormal infants. Thus, the order in which the tapes were analyzed was one week normal, one week abnormal, six months normal, six months abnormal, one year normal, one year abnormal, one week normal, etc.

The phonetician then transcribed the recordings in the International Phonetic Alphabet. A defense of this method of recording infant sounds will be found in the work of Winitz (8).

---

film. Advance payment is required. Make checks or money orders payable to Chief, Photoduplication Service, Library of Congress.

[4] Malcolm Coxe.

D. RESULTS

1. *Changes with Age*

Table 1 presents a summary of the means and standard deviations of the various phonemic elements found in the cry for the three age groups.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF VARIOUS PHONEMIC ELEMENTS IN THE CRY FOR EACH AGE GROUP

Phonemic element	One week		Six months		One year	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Total sounds per second	1.30	.58	1.51	.64	1.83	.36
Percentage of vowels	86.45	7.28	76.75	7.04	71.00	7.87
Percentage of consonants	8.10	.72	19.55	.74	24.40	.68
Number of vowels per second	1.13	.52	1.13	.40	1.28	.25
Number of consonants per second	.10	.08	.33	.26	.46	.22
Percentage of nasal sounds	20.54	18.99	31.94	20.98	26.07	16.18
Percentage of glottal/vowel	31.90	19.14	18.85	12.57	17.65	10.54

2. *Changes with Condition*

Table 2 summarizes the means and standard deviations of the phonemic elements of the cry according to condition.

Changes are noticeable with age or between conditions in certain of these elements. Each of the measures analyzed is discussed in greater detail below and a summary of the analyses of variances performed is presented in Table 3.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF VARIOUS PHONEMIC ELEMENTS IN THE CRY ACCORDING TO CONDITION

Phonemic element	Normal		Abnormal	
	Mean	*SD*	Mean	*SD*
Total sounds per second	1.74	.50	1.34	1.17
Percentage of vowels	77.63	10.31	78.50	9.21
Percentage of consonants	18.23	10.30	16.47	9.33
Number of vowels per second	1.32	.29	1.03	.47
Number of consonants per second	.35	.29	.24	.19
Percentage of nasal sounds	20.79	17.65	31.58	19.55
Percentage of glottal/vowel	21.56	11.82	23.96	19.09

TABLE 3

SUMMARY OF RESULTS OF ANALYSES OF VARIANCES FOR VARIOUS PHONEMIC ELEMENTS

Source	*df*	Total sounds	% of vowels	% of consonants	Mean number of vowels per second	Mean number of consonants per second	% of nasal sounds	% of glottal vowels
Age (A)	2	4.49*	20.00**	25.47**	1.06	15.46**	1.85	8.74**
Condition (C)	1	8.38**	.20	.85	9.21**	4.15*	4.98*	.34
A × C	2	.43	.03	.11	1.24	.16	.84	.96
Within	54							
Total	59							

* $p < .05$.
** $p < .01$.

3. *Output of Total Sounds*

The mean output of total sounds per second was calculated for each infant by dividing the total number of sounds produced by the length of the sample cry. The results of an analysis of variance performed on these data indicate that the total output of the normal infant is uniformly greater at all age levels than it is for the abnormal infant ($p < .01$). In addition, there is a significant increase with age in mean output of all sounds ($p < .05$). This is true for both the normal and the abnormal infants.

4. *Composition of Total Sound Output*

The composition of the total sounds produced was analyzed according to the percentage of vowels, diphthongs, and consonants making up the total. The results of analyses of variance performed on these data indicate that there is no significant difference between normal and abnormal infants in the percentage of specific sounds produced: however, there is a significant decrease with age in the percentage of vowel sounds produced ($p < .01$) and a significant increase with age in the percentage of consonant sounds ($p < .01$).

5. *Frequency of Specific Sounds*

The mean frequency of specific sounds was calculated for each infant by dividing the number of vowels, consonants, and diphthongs by the length of the sample cry.

a. Vowel sounds. The results of an analysis of variance indicate that normal infants tend to produce more vowel sounds per second than do abnormal infants ($p < .01$).

Table 4 shows the total number of vowel sounds that appeared most frequently according to age and condition. The sounds, identified in Figure 1, are presented in Table 4 by their corresponding number in Figure 1. The findings at age one week tend to confirm those of Irwin and Curry (5).

b. Consonant sounds. The results of an analysis of variance indicate that there is a significant increase with age in the number of consonants per second ($p < .01$). In addition, normals tend to produce more consonant sounds than do abnormals ($p < .05$); however, when the two "questionable" *S*s are excluded from this analysis, this result is no longer significant.

c. Diphthong sounds. There were no significant differences between normal and abnormal infants in the average number of diphthong sounds produced. This was true for all age groups.

TABLE 4
TOTAL NUMBER OF VOWEL SOUNDS THAT OCCURRED MOST FREQUENTLY ACCORDING TO AGE AND CONDITION

	Vowel sounds[a]							
Age	1	2	3	4	5	6	7	8
One week								
Normal	598	514	341	288	201	33	18	18
Abnormal	589	221	182	194	113	8	60	0
Six months								
Normal	541	268	204	251	137	93	4	17
Abnormal	835	248	550	240	84	106	16	31
One year								
Normal	464	164	303	145	107	49	28	10
Abnormal	527	243	399	263	112	96	32	19

[a] Vowel sounds are as follows: 1 = Schwa, 2 = Upper-low-front, 3 = Upper-low-back, 4 = Lower-mid-front, 5 = Low-back-central (stressed), 6 = Lower-high-front, 7 = Lower-low-back, and 8 = Upper-high-front. Phonetic symbols are listed in Figure 1.

6. *Variety of Different Sounds*

For each infant, the total number of different phonemes discriminable per cry sample was tabulated to determine whether any differences exist between the groups. There were no significant differences between normal and abnormal groups for all age levels.

7. *Nasality*

Since clinical observations suggested that the cries of abnormal infants were more nasal than those of normal infants, the percentage of total sounds that were nasal was calculated for each infant. The results of an analysis of variance indicate that abnormal infants show a greater percentage of nasality than do normals ($p < .05$). The greater nasalization in the abnormal infants may be a reflection of less velar control.

8. *Glottality*

Glottality, here, is simply the ejective force in initiating vowel sounds. The percentage of vowel sounds that were glottal was calculated for each infant. The results of an analysis of variance indicate that there is a significant decrease with age in the percentage of glottality ($p < .01$); however, no difference was found between normal and abnormal groups.

E. SUMMARY AND CONCLUSIONS

Tape-recorded samples of crying were obtained from three age groups (one week, six months, and one year) of brain-damaged infants and three groups

	Symbol	*Name*	*Place of tongue*	*Position of lips*	*Examples of use*
1.	ə	Schwa*	Low-central mouth area	Half open	*a*lone, sof*a*
2.	æ	Upper-low-front	Low, slightly forward	Open slightly, stretched apart	p*a*t, c*a*t
3.	U	Upper-low-back	Back, raised slightly	Half rounded, slightly protruded	b*oo*k, l*oo*k, b*u*ll
4.	ɛ	Lower-mid-front	Slightly forward and raised	Open, stretched apart	s*e*t, b*e*d
5.	∧	Low-back-central (stressed)	Low, retracted slightly	Half-open	c*u*t, p*u*tt, t*ou*ch
6.	I	Lower-high-front	Considerably forward and raised	Slightly parted and stretched	m*i*tt, s*i*t, b*ui*ld
7.	ɑ	Lower-low-back	Lowest possible, flat in mouth	Normally, widest open possible	f*a*ther, b*a*r
8.	ι	Upper-high-front	Highest and most forward possible	Slitted and most stretched apart	mach*i*ne, s*ee*, *ea*ger

* Derived from German *schwach*, meaning, in this usage, *weak* or *unstressed*.

FIGURE 1
Relevant Information on Eight Vowel Sounds

of normal infants matched for age and sex. These were then analyzed for phonetic content and the following results were obtained:

1. The phonetic content of the cry for both normal and abnormal infants is predominantly made up of vowel sounds. These tend to decrease with age as the number of consonant sounds increases.
2. There is no significant difference between normal and abnormal infants in the percentage of specific sounds produced.
3. Normal infants produce significantly more total sounds than do abnormal infants; they also produced more vowel sounds than do abnormal infants.
4. Abnormal infants produce significantly more nasal sounds than do normal infants.

REFERENCES

1. FISICHELLI, R. M. A study of prelinguistic speech development of institutionalized infants. Unpublished Ph.D. dissertation, Fordham University, New York, 1950.
2. FISICHELLI, V. R., & KARELITZ, S. The cry latencies of normal infants and those with brain damage. *J. Pediat.,* 1963, **62**, 724-734.
3. IRWIN, O. C. Speech and sound elements during the first years of life: A review of the literature. *J. Speech Disorders,* 1943, **8**, 109-121.
4. IRWIN, O. C., & CHEN, H. P. A reliability study of speech sounds observed in the crying of newborn infants. *Child Devel.,* 1941, **12**, 351-368.
5. IRWIN, O. C., & CURRY, T. Vowel elements in the crying vocalizations of infants under ten days of age. *Child Devel.,* 1941, **12**, 99-109.
6. KARELITZ, S., & FISICHELLI, V. R. The cry thresholds of normal infants and those with brain damage. *J. Pediat.,* 1962, **61**, 679-685.
7. KARELITZ, S., KARELITZ, R., & ROSENFELD, L. S. Infants' vocalizations and their significance. In P. W. Bowman & H. V. Mautner (Eds.), *Mental Retardation: Proceedings of the First International Medical Conference.* New York: Grune and Stratton, 1960. Pp. 439-446.
8. WINITZ, H. Spectrographic investigation of infant vowels. *J. Genet. Psychol.,* 1960, **96**, 171-181.

Department of Pediatrics
The Long Island Jewish Hospital
270-05 76th Avenue
New Hyde Park, Long Island, N. Y.

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 127-133.

THE EFFECT OF SHORT-TERM SENSORY DEPRIVATION ON ROTE LEARNING*[1]

Department of Psychology and Industrial Management, the University of Tennessee

CABOT L. JAFFEE

A. INTRODUCTION

Reviews on the effects of sensory deprivation have been done by Lilly (12), Solomon, Leiderman, Mendelson, and Wexler (16), Solomon, Leiderman, Trumbull, Kubzansky, Mendelson, and Wexler (17), Fiske (6), and Brownfield (3). Each of these reviews reports that subjects remaining for long periods of time under conditions of restricted stimulation exhibit behavioral changes very different from subjects under conditions of usual stimulation.

However, there is also a great deal of evidence that indicates that behavioral changes may occur after relatively brief periods of sensory deprivation. Hochberg, Triebel, and Seaman (10) found that hallucinations may appear within 20 minutes. Ormiston (13), depriving his subjects by means of translucent goggles and earmuffs, found that only 30 minutes of deprivation were needed to produce changes in threshold for apparent movement. Rosenbaum, Dobie, and Cohen (14) found that increments in visual efficiency would result from a five-minute period of deprivation. Visual efficiency was measured by tachistoscopic presentation of words and recording of recognitive thresholds.

This evidence seems to indicate that periods of time as brief as five minutes may produce increments in performance on a variety of tasks.

Rote memory learning has been extensively investigated relative to the effects of sensory deprivation. There have been two studies that have found increments in performance on a rote learning task after periods of deprivation (4, 19), while a number of studies have found decrements or no changes on rote learning after deprivation (1, 2, 5, 7, 8, 9, 15, 20, 21).

The present research was designed to explore the short-term effects of visual and auditory deprivation on the memorizing of a list of words. This study also compared the performance of groups under two periods of deprivation.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on July 25, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study is based upon a dissertation submitted to the graduate school of The Florida State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

B. METHOD

1. *Subjects*

The subjects were 180 public school pupils from the Gadsden County, Florida, school system. They were selected from the sixth grade and grouped into two levels of intelligence: high, 113-143; and low, 81-101. The Otis Quick Scoring Mental Ability Test was used to determine intelligence quotients of the subjects. On the basis of individuals' performance on the Otis, all groups received an equal number of high *IQ* and low *IQ* pupils. The pupils were divided into six groups with 30 individuals in each. One group received a five-minute deprivation period between performance measures; the second received a one-minute period of deprivation. Another received a five-minute time lapse between performance measures, but no deprivation, while the fourth received a one-minute time lapse. The fifth and sixth groups had five and one minutes, respectively, of extra time to work on the second word list before a performance measure was taken.

Care was taken to insure a minimum of contact between experimental subjects. A group of subjects was removed from a single classroom, and upon each child's completion of the task he or she was sent directly back to the classroom without having any contact with the waiting group.

2. *Apparatus*

The apparatus consisted of the following components that were housed in a small room that contained a table and two straight-back chairs as well as the experimental apparatus.

1. Deprivation material: A pair of padded earphones used to diminish auditory stimulation, and a pair of blacked-out rubber goggles for total visual deprivation.
2. Auditory stimulus apparatus: A tape recorder that presented to the subject, through a set of earphones, the predetermined list of stimulus words.
3. Visual stimulus apparatus: A slide projector that presented on a screen before the subject the predetermined list of stimulus words.

The auditory and visual stimuli were synchronized so as to have the stimulus words presented simultaneously through both media. In this way maximum stimulation was effected.

4. Subject's viewing screen: A 24- by 30-inch flat white screen hung at eye level four feet from the subject. The subject sat in a chair alongside the table upon which the tape recorder and slide projector were situated. The experimenter sat behind and to the subject's right, directly behind the slide projector. The subject faced the screen.

5. Visual stimulus slides: The stimulus material utilized for the experiment was a set of 2- by 2-inch slides. Forty such slides were constructed with different words on each. Two lists of 20 words each were made up using words having similar frequencies of occurrence in general reading materials (18). The stimulus for the interval group was a series of geometric forms described by Kennedy *et al.* (11).

6. Auditory stimulus: The stimulus material utilized here was the tape recording of these same 40 words presented to the subject by means of earphones.

A trial consisted of the presentation of one list of stimulus words. All subjects received the same stimuli arranged in the same sequence.

3. *Procedure*

The procedure, except for the experimental treatment (deprivation, time interval, or work) which was interpolated between Trial 1 and Trial 2, was identical for all the subjects. Both 20-word lists were presented to each subject. All subjects were brought into the experimental room and seated facing the screen. They were then told to try to remember as many of the words as they were able from the list of 20 words presented.

As previously stated, a trial consisted of presenting 20 ordered words to the subject and having him repeat orally as many as possible. At the conclusion of Trial 1 and before Trial 2, the experimental condition was interpolated. This was either a five- or one-minute time lapse (deprivation period), or a study interval, during which a group was given the opportunity to work for an additional period of time on the second list. The subjects in the five-minute work group had the second list repeated five times and the subjects in the one-minute work group had the second list repeated twice, while only a single presentation of each stimulus word was given to the other treatment groups. All subjects were then administered the second list and tested in an identical manner.

C. RESULTS

The dependent measure used in this study was the number of words recalled from a list of 20 words presented. Since the major interest of this study was to examine the effects of the different conditions, measured over trials, the interaction between these variables and the time and intelligence variables assumes major importance in the results. A 3 × 2 repeated measure analysis of variance indicated two significant main effects between subjects. The F ratios between conditions ($F = 9.01$, 2 and 168 *df*) and intelligence levels

($F = 39.14$, 1 and 168 *df*) were significant beyond the .01 level. The work condition group memorized a greater mean number of words than did the deprivation and the interval group, as indicated by a Duncan Multiple Range Test. The intelligence of the subjects was related positively to the intelligence level, in that the high intelligence group memorized a greater mean number of words than did the low intelligence group.

In terms of the within subjects analysis, a number of comparisons were significant. The F ratios between the means of Trial 1 and Trial 2 were significant at less than the .01 level ($F = 30.90$, 1 and 168 *df*), resulting from an overall increase in the means over trials. The interaction of conditions by trials ($F = 20.35$, 2 and 168 *df*) and time by trials ($F = 8.68$, 1, 168 *df*) were both significant at less than the .01 level. A Duncan Multiple Range Test revealed that the significant conditions by trials interaction was apparently the result of the fact that both the deprivation and work groups significantly increased the number of words memorized over trials, while the interval group showed no such increase.

The time by trials interaction is the result of a significantly greater mean number of words memorized in the longer time group, from Trial 1 to Trial 2.

Thus, with the exception of the interval group, all conditions produced a statistically significant increase in mean number of words memorized from Trial 1 to Trial 2. That is, only the subjects in the interval group remained stable or decreased from first to second trial. Deprivation had the same effect as increasing the time allotted for work, in that both these groups increased significantly in mean number of words remembered from Trial 1 to Trial 2.

In summary, deprivation resulted in an increase in mean number of words memorized, while an interval of time between trials decreased or failed to affect the number of words remembered from Trial 1 to Trial 2. The longer time period over trials resulted in a statistically significant increase in mean number of words recalled, while the shorter time period groups showed no significant differences between trials.

In terms of main effects, the work groups recalled the most words, deprivation and interval groups having significantly lower mean scores. Intelligence proved significant in that, for overall number of words memorized, subjects in the higher intelligence group did significantly better than those in the low group. However, when intelligence was analyzed by trials, there was no significant interaction found.

On overall performance there were no differences found between the time groups, but when this variable was analyzed by trials, a significant difference was found. The differences appeared to lie in the work and deprivation groups rather than with the interval group.

D. DISCUSSION

In terms of the conditions established for this study, several significant relationships were observed.

Overall, subjects at both intelligence levels responded alike to a given experimental condition. The conditions of deprivation and work increased the number of words recalled, while a time interval tended to decrease recall from Trial 1 to Trial 2. The variability about this general formulation provides some of the most interesting findings in the study. For the deprivation groups the greatest changes took place for the longer time groups. These findings are similar to those found by Rosenbaum, Dobie, and Cohen (14), where a five-minute period of deprivation facilitated perceptual functioning. The present study provides a means of evaluating the effects of deprivation as opposed to a five-minute performance enhancing set, which they hypothesized as a possible cause for the increment in performance. In the present study the interval group showed no such facilitation of performance from Trial 1 to Trial 2. In fact, a slight decrement resulted in some cases.

In comparing the data for the work and deprivation groups, it was found that a group presented with the same list twice did as well as a group given a five-minute period of deprivation before being presented with the list once. This finding is one of major interest to the present study, as this comparison provides a means of evaluating the relative facilitation in word memorization caused by a five-minute period of deprivation.

Probably the most striking finding was that the effect of a five-minute period of deprivation produced an increase equal in magnitude to that produced by a group presented with a list of words for double the period of time. The study grew out of the observation that, on the basis of the existing literature on the subject, the effects of deprivation on cognitive and perceptual functioning seemed to be reliant on time factors as well as situational variables. The results of this study indicate that under the conditions of a five-minute deprivation period the results appear consistent and congruent with previous literature, in that a five-minute time period and minimal stimulation appear to enhance cognitive functioning to a greater degree than do longer times and conditions of patterned deprivation.

The results of this study by no means answer fully any questions with respect to varying periods of deprivation other than the one- and five-minute periods used in this experiment. They do, however, indicate strongly that the differences exist due to deprivation rather than to any other situational variables. This may offer a point of reference from which to work in approaching the question more fully.

In any event, these data allow the logical conclusions that in situations involving cognitive tasks, and using sixth-grade children, with words memorized as a dependent measure, the results obtained should hold. That is, a five-minute period of deprivation results in an increment in performance, while no such increment occurs in a group not so treated.

E. SUMMARY

This present study was aimed at evaluating the effects of three treatments—visual and auditory deprivation, a time interval, or additional work on the number of words to be memorized—by high and low intelligence subjects given varying times of deprivation, interval, or work.

One hundred and eighty subjects were used. These were subdivided equally over all levels of the variables: condition, intelligence, and time.

A five-minute period of deprivation was found to increase performance, which was unaffected by a one-minute period of deprivation. A time interval elapsing between performance measures had no effect on the number of words memorized by subjects on the second list of words.

REFERENCES

1. ARNHOFF, F. W., LEON, H. V., & BROWNFIELD, C. A. Sensory deprivation: On human learning. *Science,* 1962, **138**, 899-900.
2. BEXTON, W. H., HERON, W., & SCOTT, T. H. Effects of decreased variation in the sensory environment. *Can. J. Psychol.,* 1954, **8**, 70-76.
3. BROWNFIELD, C. A. Deterioration and facilitation hypotheses in sensory-deprivation research. *Psychol. Bull.,* 1964, **61** (No. 4), 304-313.
4. COHEN, S. I., SILVERMAN, A. J., BRESSLER, B., & SHMAVONIAN, B. Problems in isolation studies. In P. Solomon *et al.* (Eds.), *Sensory Deprivation.* Cambridge: Harvard Univ. Press, 1961. Pp. 114-129.
5. DOANE, B. K., MAHATOO, W., HERON, W., & SCOTT, T. H. Changes in perceptual function after isolation. *Can. J. Psychol.,* 1959, **13**, 210-219.
6. FISKE, D. W. Effects of monotonous and restricted stimulation. In D. Fiske & S. Maddi (Eds.), *Functions of Varied Experience.* Homewood, Ill.: Dorsey Press, 1961. Pp. 106-144.
7. HERON, W. The pathology of boredom. *Sci. Amer.,* 1957, **196** (1), 52-56.
8. HERON, W., BEXTON, W. H., & HEBB, D. O. Cognitive effects of a decreased variation in the sensory environment. *Amer. Psychol.,* 1953, **8**, 366 (Abstract).
9. HERON, W., DOANE, B. K., & SCOTT, T. H. Visual disturbances after prolonged perceptual isolation. *Can. J. Psychol.,* 1956, **10**, 13-18.
10. HOCHBERG, J. E., TRIEBEL, W., & SEAMAN, G. Color adaptation under conditions of homogeneous visual stimulation (Ganzfeld). *J. Exper. Psychol.,* 1951, **41**, 153-159.
11. KENNEDY, W. A., TURNER, A. J., & LINDNER, R. Effectiveness of praise and blame as a function of intelligence. *Percept. & Motor Skills,* 1962, **15**, 143-149.

12. LILLY, J. C. Mental effects of reduction of ordinary levels of physical stimuli on intact, healthy persons. *Psychiat. Res. Rep.*, 1956, **5**, 1-9.
13. ORMISTON, D. W. The effects of sensory deprivation and sensory bombardment on apparent movement thresholds. Unpublished Doctoral dissertation, Purdue University, Lafayette, Indiana, 1958. *Diss. Abst.*, 1958, **18**, 2200-2201.
14. ROSENBAUM, G., DOBIE, SHIRLEY, I., & COHEN, B. D. Visual recognition thresholds following sensory deprivation. *Amer. J. Psychol.*, 1959, **72**, 429-432.
15. SCOTT, T. H., BEXTON, W. H., HERON, W., & DOANE, B. K. Cognitive effects of perceptual isolation. *Can. J. Psychol.*, 1959, **13**, 200-209.
16. SOLOMON, P., LEIDERMAN, P. H., MENDELSON, J., & WEXLER, D. Sensory deprivation: A review. *Amer. J. Psychol.*, 1959, **114**, 357-363.
17. SOLOMON, P., KULZANSKY, P. E., LEIDERMAN, P. H., MENDELSON, J. H., TRUMBULL, R., & WEXLER, D. (Eds.). Sensory Deprivation. Cambridge: Harvard Univ. Press, 1961.
18. THORNDIKE, E. L., & LORGE, I. The Teacher's Word Book of 30,000 Words. New York: Bur. of Publ., Teachers Coll., Columbia University, 1944.
19. VERNON, J. A., & HOFFMAN, J. Effects of sensory deprivation on learning rate in human beings. *Science*, 1956, **123**, 1074-1075.
20. VERNON, J. A., & MCGILL, T. E. The effect of sensory deprivation on rote learning. *Amer. J. Psychol.*, 1957, **70**, 637-639.
21. ZUBEK, J. P., AFTANAS, M., HOSEK, J., SANSOM, W., SCHLEIDERMAN, E., WILGOSH, L., & WINOCUR, G. Intellectual and perceptual changes during prolonged perceptual deprivation: low illumination and noise level. *Percept. & Motor Skills*, 1962, **15**, 171-198.

Department of Psychology and Industrial Management
University of Tennessee
Knoxville, Tennessee

$30.00 per annum
$15.00 per volume
$7.50 per half volume

BIMONTHLY
Three volumes per year
Immediate publication

November, 1966
Volume 64, Second Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

NOVEMBER, 1966

(Manuscripts are printed in the order of final acceptance)

Racial stereotypes in young children 137
By Raymond G. Taylor, Jr.

Psychometric properties of the Wiggins and Winder Peer Nomination Inventory 143
By Marvin Siegelman

Methodology of relative sugar preferences in laboratory rats and deer mice 151
By Mahlon W. Wagner and J. Timothy Rowntree

Factors and factor analyses 159
By J. A. Gengerelli

Just how far can social change change personality? 167
By Ralph Mason Dreger

(over)

Copyright © 1966 by The Journal Press
Provincetown, Massachusetts, U. S. A.
All rights reserved

Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts

A study of form discrimination in the opossum 193
By W. T. James and John McFarland

The effects of anxiety, stress, and task difficulty on performance on a verbal learning task 199
By Donald W. Brown

Giannell's criminosynthesis theory applied to female homosexuality . . . 213
By A. Steven Giannell

The effect of DNA on operant conditioning 223
By C. Beaulieu and L. Solyom

Comparative birth weights of schizophrenics and their siblings . . . 227
By Ellen A. Lane and George W. Albee

Brain blood-shift theory: Intersensory rivalry, time error, and closure illusions 233
By Conrad Chyatte

The study of target size in production of brightness, hue, and saturation for intermittent photic stimulation in the red end of the visible spectrum . 241
By S. Howard Bartley and Richard J. Ball

Sex differences in smoking attitude, behavior, and beliefs of college students 249
By Daniel J. Baer

The effects of phenothiazine drugs on verbal conditioning of schizophrenics 257
By Norris Vestre

Male-female perceptions of the female role in the United States . . . 265
By Anne Steinmann and David J. Fox

Desensitization factors in the hypnotic treatment of phobias 277
By Joseph R. Cautela

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 137-142.

RACIAL STEREOTYPES IN YOUNG CHILDREN*

The Institute for the Achievement of Human Potential, Philadelphia

RAYMOND G. TAYLOR, JR.

A. INTRODUCTION

It is well-known that negatively valued stereotypes about the Negro race are common in areas where the racial attitudes of whites predominate, but there is limited knowledge as to how greatly negatively valued stereotypes about their own race are shared by Negroes themselves. The present study was designed to learn to what extent young children absorb commonly held negatively valued stereotypes of the Negro race and to apply a modern multivariate technique to the solution of the problem.

B. PREVIOUS RESEARCH

The investigations of Mary E. Goodman (2) showed that stereotypes of this kind begin to be absorbed early in childhood. She concluded from her studies that a child of 3 is able to develop a consciousness of self that is inseparable from a consciousness of others. By 4 years of age the child's concept of other persons is well enough established for him to begin to make rational comparisons, attaching values to the differences he establishes. This means that at age 4 the child is capable of prejudging an individual on the basis of his membership in a generalized group, especially race. This tendency to classify and generalize becomes greater as the child passes through his elementary school years. Although the step from classification to establishing value judgments is not a necessary one, normally the older the child becomes, the more he assigns values to group membership that correspond to those commonly found in the adult world. Thus, along with the ability to differentiate, the child learns the specific stereotypes applied by adults to members of particular groups, especially race.

Knowledge of this kind has been used as a basis for previous studies dealing with racial attitudes in children, especially those of Horowitz (3) and Blake and Dennis (1). In these studies it was recognized that numerous variables effect racial attitudes. However, it was not until about 1950 that social scientists possessed the analytical means of studying the effects of several

---

* Received in the Editorial Office, Provincetown, Massachusetts, on July 27, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

variables simultaneously and, consequently, the results of earlier research were based largely on univariate analyses. The present study utilizes a four-way analysis of variance, a multivariate procedure that reveals the effects of single variables while controlling for the effects of the others. The hypotheses to be tested appear in the Results section.

C. SAMPLE

It was decided to study the attitudes of a group of children who were diverse with respect to age, race, intelligence, and socioeconomic background. The sample was chosen in cooperation with public school principals in various areas of the city of Philadelphia, each of whom volunteered several classes of pupils based on the sampling criteria. The following information was made available about each child in the sample: sex, age, race, grade in school, intellectual ability, and socioeconomic status of the parents. Only children in the first and third grades of school were chosen in order to keep the mean age low and to minimize the overlapping of ages within the grade stratification. In most Philadelphia public schools the children are divided into three ability groups for each elementary grade. For this study children were chosen from the highest and lowest groups to minimize the overlapping of intellectual ability. The socioeconomic status of each child was determined by the teacher and the school principal. All pupils who had at least one parent who graduated from college or whose parents had a combined income of at least $12,000 per year were placed in socioeconomic group A. The pupils whose parents were on relief or whose combined income was less than $4,000 per year were placed in group C. All others were placed in group B.

A total of 229 children were obtained. These children were reasonably well distributed within the stratifications of sex, race, grade in school, location, ability group, and socioeconomic status. However, it was impossible to obtain first grade Negro children in the lowest ability group because these children had so little experience with test taking that they could not complete the test in the amount of time allowed for the administration. It was also impossible to obtain a meaningfully large sample of Negro children in socioeconomic group A.

D. DESIGN OF THE INSTRUMENT

The study required an instrument that would measure the existence and the extent of negatively valued racial stereotypes in elementary school pupils. In a well designed experiment by Horowitz (3), pictures were used to determine the value judgments of children. He asked the children (*a*) to rank portraits of individuals in order of preference, (*b*) to select from portraits individuals whom they would like for companions, and (*c*) to decide whether they would like to participate in social situations that were portrayed in pictures.

The use of pictures for testing has at least three distinct advantages: (*a*) all school children are accustomed to hearing stories based on pictures, (*b*) pictures help to hold the attention of children, and (*c*), as used in this study, pictures provide an objective method of scoring.

Nine photographs, each of which showed a Negro and a white person involved in the same activity, were reproduced and given to each child (see Figure 1). Anecdotes were written to accompany these photographs, each implying that a commonly held, negatively valued stereotype was true of one of the two individuals in the picture. Each child was asked to identify the person to whom the stereotype applied. For example, as the children looked at the seventh picture, the following anecdote was told:

> Now look at Picture 7. These two boys are good friends; they live in the same neighborhood and they go to school together. One is very smart and gets good grades; the other one, I am sorry to say, is sort of dumb. Which one do you think is smart? Point to the smart one.

As decisions were made for each picture, the writer scored the answers on a separate answer sheet. To standardize the procedure and assure uniformity all administrations were done completely by the writer.

The nine stereotypes, taken mostly from Goodman (2, p. 104 *f.*) were (*a*) lazy, (*b*) dirty, (*c*) dishonest, (*d*) careless, (*e*) unfitting as neighbor, (*f*) incompetent, (*g*) dumb, (*h*) bad, and (*i*) not nice.

The instrument was scored on the basis of how many times a negatively valued stereotype was associated with a Negro subject. Thus, a score of 8 would mean that on eight pictures the negatively valued stereotype was associated with the Negro child, and on one picture with the white child.

The instrument possesses at least two limitations. Some photographs indicate obvious differences in the race of the subjects, while others are more subtle; and yet each picture was given the same weight when scoring. Some pictures contained considerable objective evidence that, if used, would influence the response. However, the pictures that contained such information were designed to give the Negro subjects the positively valued response in order to prevent an unrealistic elevation in the score.

E. RESULTS

Hypothesis 1: Children of elementary school age possess negatively valued stereotypes of Negroes. This hypothesis was tested on the basis of the total distribution of scores without reference to the stratifications within the

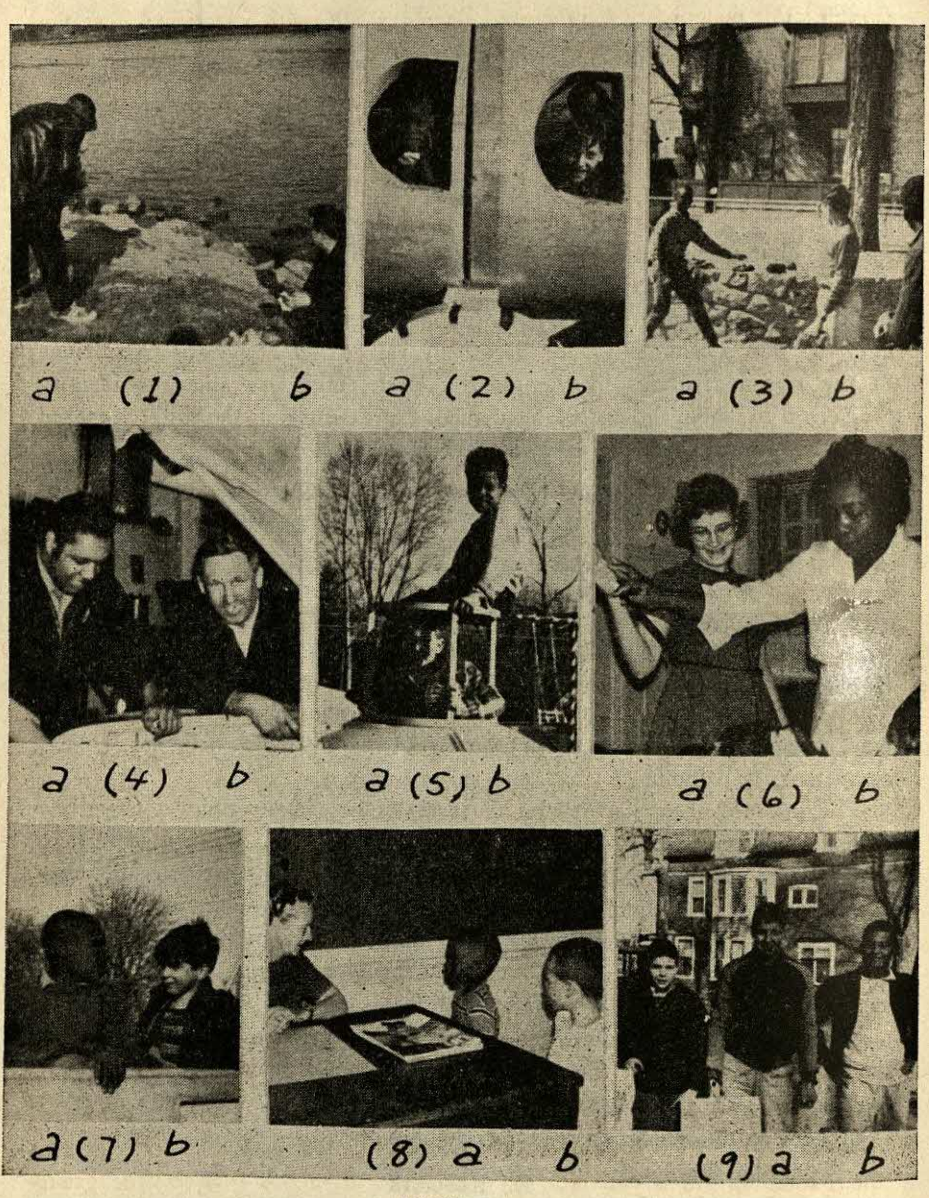

FIGURE 1
A Pictorial Test Designed to Measure Certain Social Attitudes Among Children

sample. The mean score for the total sample is 5.06 with a standard deviation of 1.53. If one assumes that chance would be at the midpoint of the scale where whites and Negroes receive an equal number of negatively valued responses, the observed mean is only slightly above chance expectancy ($p < .10$).

The remaining hypotheses involve the principal analytical work of the study: that is, the measurement of the effect of each of several variables while controlling for the effect of the others. This was accomplished by arranging the results in a five dimensional table and performing a four-way analysis of variance (see Table 1).

TABLE 1
ANALYSIS OF VARIANCE

Source	Sum of squares	*df*	*F* ratio	*p*
Race	12.0	1	6.30	$< .025$
Grade	77.0	1	40.50	$< .005$
Ability	0.0	1	0.00	
Socioeconomic status	28.0	2	7.40	$< .005$
Interactions combined	27.0	15	.95	
Within	394.0	208		
Total	538.0	228		

Hypothesis 2: Race itself effects racial attitudes in children. The above analysis indicates that the effect of race is significant ($p < .025$) after correcting for the effects of the remaining main variables. The mean score of the 141 white children was 5.2; of the 88 Negro children, 4.7. Both means fell above the "nonprejudiced" score of 4.5, although the Negro mean was well within chance expectancy. It is interesting to note that every Negro child in the sample chose the white boy in Picture 2 as the "clean" one. Likewise, only six Negro children indicated that they thought the Negro boy in Picture 7 was the "smart" one.

Hypothesis 3: Socioeconomic status effects racial attitudes in children. Socioeconomic status was found to have a significant effect on the test scores ($p < .005$), independent of the other main variables. The mean score of the 31 children in group A was 5.6; of the 120 children in group B, 5.2; and of the 78 children in group C, 4.6.

Hypothesis 4: Intelligence effects racial attitudes in children. The mean score of both ability groups was 5.1. Even after correcting for the effect of the remaining three variables, no significant difference between the groups was indicated.

Hypothesis 5: Grade in school effects racial attitudes in children. Extremely significant differences between children in grades one and three were indicated after controlling for the effects of the other main variables ($p < .005$). The mean of 96 children in grade one was 5.7; of the 133 children in grade three, 4.6.

Two criticisms might be made of this analysis: (*a*) It is possible that children in grade three were better able than children in grade one to control their answers and appear "less prejudiced." If this happened, however, we are left to wonder why it did not likewise influence the scores of the "fast" ability group. (*b*) The selection of only two grades makes it difficult to generalize accurately to the whole range of elementary school grades. The results apply best to children between ages 6 and 10.

F. SUMMARY

This study of 229 elementary school children in the city of Philadelphia in the spring of 1965 indicated the following:

1. Children between the ages of 6 and 10 possess negatively valued stereotypes of Negroes.
2. Controlling for the effects of ability, socioeconomic status, and grade in school, white children hold more of these stereotypes than do Negro children. Negro children, however, accept to a very high degree the stereotypes of "dirty" and "dumb" as applied to their own race.
3. Controlling for the effects of race, ability, and grade in school, the higher the socioeconomic station of the child the more likely he is to possess such stereotypes of Negroes.
4. Intelligence as determined by ability grouping in school does not effect a child's racial stereotypes.
5. Controlling for the effects of race, socioeconomic status, and ability, the older the child the more likely he possesses negatively valued stereotypes of Negroes.

REFERENCES

1. BLAKE, R., & DENNIS, W. The development of stereotypes concerning the Negro. *J. Abn. & Soc. Psychol.,* 1943, **38**, 525-531.
2. GOODMAN, M. E. Race Awareness in Young Children. Cambridge, Mass.: Addison-Wesley Press, 1952.
3. HOROWITZ, E. L. The development of attitude toward the Negro. *Arch. Psychol.,* 1936, **28**, No. 194.

8801 Stenton Avenue
Philadelphia 18, Pennsylvania

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 143-149.

PSYCHOMETRIC PROPERTIES OF THE WIGGINS AND WINDER PEER NOMINATION INVENTORY*[1]

Department of Social and Psychological Foundations, City College of New York

MARVIN SIEGELMAN

A. INTRODUCTION

In 1961 Wiggins and Winder (8) reported a sociometric measure of adjustment for elementary school boys called The Peer Nomination Inventory (PNI). Most peer nomination techniques give only an index of popularity (1), but the PNI measures several important personality dimensions as well. Recent studies (6, 9) have supported the reliability and validity of the PNI. The purpose of the present investigation is to report a factor analysis of the PNI items and to compare these results with the Wiggins and Winder factor analysis of the PNI items (8). Greater insight into the structure and meaning of the PNI may be gained in this way. Reliabilities and correlations among the factors will also be noted.

B. METHOD

1. *Subjects*

A middle-class group of white fourth-, fifth-, and sixth-grade boys ($N =$ 113) comprised the sample. All the boys attended an elementary school in the Bronx, New York, and were predominantly Jewish. The PNI was group administered during one class period and scored in the manner described by Wiggins and Winder (8, p. 660). Each boy in a given class rated all the other boys in his class. The standardization sample of 710 boys noted by Wiggins and Winder were also fourth-, fifth-, and sixth-graders from a middle-class community.

2. *The Peer Nomination Inventory*

The PNI consists of 62 statements including four beginning filler items used only for demonstration purposes. Five major adjustment areas covered by 56 items presented in random order are Aggression (12 items), Dependency

---

* Received in the Editorial Office, Provincetown, Massachusetts, on August 12, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study was supported by a grant from the City College Research Foundation. The author is grateful to Diane Delfino for her conscientious assistance.

(12 items), Withdrawal (12 items), Depression (12 items), and Likeability (eight items). The two remaining items relate to socioeconomic status. Each child indicates which children in his class fit the description of a given item. The attempt was made in the present study to duplicate closely the PNI form and administration used by Wiggins and Winder. A detailed description of the development, use, and administration of the PNI is presented in the Wiggins and Winder (8) monograph.

C. Results and Discussion

A product-moment correlation matrix for the 58 PNI items was computed. These coefficients were based on continuously scored percentage of class nominations for a given boy (8, p. 660). A principal component factor analysis (3), using ones in the diagonal as communality estimates (5), and a subsequent varimax rotation (4) of the 58 × 58 correlation matrix was then calculated. Because the rotated factor loadings would be applied to continuously scored percentage of class ratings to get factor scores, these ratings were used to compute the original product-moment correlation matrix. The above procedure differed from the Wiggins and Winder (8) use of tetrachoric correlations based on dichotomously scored PNI items. In addition, Wiggins and Winder reported a factor analysis of 48 items covering aggression, dependency, withdrawal, and depression, while the present factor analysis also included the eight likeability items and the two socioeconomic items.

Table 1 presents a four factor rotation of the 58 PNI items. Five, six, and seven factor rotations were also calculated and considered. The four factor solution was selected for presentation in that it accounts for 65.4 per cent of the total factor variance, it depicts the most clear-cut and meaningful factors, and it does not differ appreciably from the five, six, and seven factor rotations. The original PNI variable scales reported by Wiggins and Winder (8) before they factor-analyzed the items included items 1-12 for Aggression, 13-24 for Dependency, 25-36 for Withdrawal, 37-48 for Depression, 49-56 for Likeability, and 57-58 for Socioeconomic Status.

Factor I in Table 1 is primarily composed of aggression *and* dependency items. Most of the highly saturated items imply a tendency to attract attention by means of deviant classroom behavior. Factor I is designated Aggression-Dependency. A close association between aggression and dependency was also reported by Bandura (2) for preadolescent aggressive boys who displayed a great deal of dependency behavior toward their teachers, parents, and peers. Wiggins and Winder (8) reported a Hostility Factor II, in their PNI item factor analysis, which accounted for 18.2 per cent of the total variance, and

TABLE 1

Rotated Factor Loadings, Communalities, and Percentage of Variance

PNI item	I	II	III	IV	h^{2b}
Aggression[a]					
1. He's always acting up.	.82	.07	.04	.12	.69
2. When he doesn't get his way, he gets real mad.	.70	.03	—.02	.27	.56
3. If someone gets in his way, he shoves them out of the way.	.80	—.10	—.21	.00	.68
4. He's just plain mean.	.67	—.02	—.31	.04	.55
5. He's really wild.	.74	.16	.11	.01	.59
6. He makes fun of people.	.80	.02	—.16	—.02	.66
7. He doesn't pay attention to the teacher.	.80	.37	—.15	—.10	.82
8. He seems to have a chip on his shoulder.	.66	.35	—.15	.13	.60
9. He tries to get other people in trouble.	.89	.09	—.16	—.02	.83
10. He always messes around and gets in trouble.	.87	.11	—.08	—.06	.78
11. He says he can beat everybody up.	.68	—.08	—.01	—.01	.47
12. He likes to pick on little kids.	.78	—.01	—.05	—.02	.61
Dependency					
13. He likes an audience all the time.	.71	.03	.32	.01	.62
14. He feels a lack of attention.	.46	.71	—.09	.00	.72
15. He needs attention very badly.	.63	.57	—.10	.05	.73
16. He talks to the teacher all the time.	.48	.36	.11	.27	.44
17. He's trying hard to get popular.	.51	.16	.17	.00	.32
18. He's always playing the clown and wants everyone to laugh at him.	.78	.09	.06	.00	.62
19. He wants everything done for him.	.77	.30	—.10	.03	.69
20. He wants to show off in front of the kids.	.84	.01	.06	—.10	.72
21. He'll talk out to get attention.	.87	.04	.01	—.05	.77
22. He just acts sort of babyish all the time.	.45	.40	—.01	.36	.62
23. He tries to get attention.	.87	.12	.06	—.01	.78
24. He acts as if he's sort of a baby.	.49	.34	—.07	.44	.56
Withdrawal					
25. He will always play by himself.	—.03	.89	—.10	.02	.80
26. No one plays with him.	.02	.87	—.11	—.04	.77
27. He's not interested in anything.	.49	.58	—.21	—.12	.63
28. He doesn't play.	—.14	.85	.00	.10	.75
29. He is sort of ignored.	.15	.85	—.14	—.00	.77
30. Hardly any boys like to play with him.	.10	.84	—.17	.01	.74
31. He doesn't want to play.	—.11	.87	—.09	.07	.78

TABLE 1 (*continued*)

PNI item	I	II	III	IV	h^{2b}
32. He's the last person picked.	.17	.78	—.10	.05	.65
33. On the playground he just stands around.	—.19	.60	—.08	.21	.45
34. He just seems sort of lost.	.24	.83	—.11	.08	.77
35. He doesn't have very many friends.	.08	.84	—.12	.12	.74
36. He doesn't care to join in and play on the playground.	—.06	.80	—.13	.21	.70
Depression					
37. Someone makes fun of him and he starts crying.	—.03	.08	—.17	.83	.72
38. He cries if you hurt his feelings.	—.05	—.01	—.20	.86	.78
39. He cries when he doesn't do something right.	—.07	.02	—.07	.86	.76
40. No matter what he does, it's wrong.	.63	.40	—.18	—.10	.60
41. He's a little too sensitive to everybody.	.16	.22	—.07	.66	.65
42. He feels left out.	.13	.86	—.05	.08	.76
43. He's not sure of himself in anything.	.50	.61	—.11	—.00	.63
44. He says he can't do things.	.18	.69	—.14	.15	.55
45. He cries when he doesn't know how to play.	.02	.22	—.03	.68	.51
46. He just can't stand anybody laughing at him.	.70	.13	—.23	.26	.62
47. He's sort of unhappy.	.19	.72	—.22	.12	.62
48. He seems to think that he's nobody.	.17	.87	—.19	.06	.82
Likeability					
49. He has a lot of friends.	.04	—.33	.73	.02	.65
50. He's one of the kids I like.	—.05	—.23	.85	—.08	.79
51. He's the kind of kid I like.	.00	—.18	.85	—.12	.77
52. He's a good friend of mine.	—.10	—.20	.75	.04	.61
53. I'm one of his friends.	—.02	—.11	.87	—.12	.78
54. All the kids like him.	.11	—.16	.77	—.12	.65
55. He's a good sport.	—.16	—.09	.80	—.19	.70
56. He's a nice guy.	—.11	—.06	.86	—.10	.78
Socioeconomic Status					
57. He's got a real nice bike.	.32	—.12	.27	.06	.20
58. He has a lot of money to spend.	.40	—.08	.13	—.02	.19
Per cent common variance	.39	.34	.17	.10	
Per cent total variance	.25	.22	.11	.07	

Note: I = Aggression-Dependency factor, II = Withdrawal-Depression factor, III = Likeability factor, and IV = Crying factor.

[a] Original PNI scale titles used by Wiggins and Winder (8).

[b] Communalities.

a separate Attention-getting Factor V which accounted for 10.2 per cent of the total variance. These Factors II and V were hardly unrelated to each other, however, in that the correlation between them was .743. In a factor analysis of the original five PNI scales, Wiggins and Winder (8) also indicated a distinct Aggression and Dependency Factor with factor loadings of .84 for Aggression and .85 for Dependency.

Factor II in Table 1 has its heaviest saturations on withdrawal and depression items, and is very similar to the Wiggins and Winder Social Isolation Factor I that accounted for 27.1 per cent of the total variance. Factor II in Table 1 is labelled Withdrawal-Depression. Several of the original depression items constitute the independent Factor IV, Crying, in Table 1. Except for item 41, the same items were highly saturated and formed an independent Factor II in the Wiggins and Winder data, and there comprised 6.3 per cent of the total variance.

TABLE 2
CORRELATIONS BETWEEN FOUR PNI FACTORS AND RELIABILITY ESTIMATES*

PNI factor	I Aggression-Dependency	II Withdrawal-Depression	III Likeability	IV Crying
I Aggression-Dependency	.960	.368	—.110	.059
II Withdrawal-Depression		.969	—.353	.214
III Likeability			.939	—.245
IV Crying				.864

* PNI factors derived from data in present study. Generalized Kuder-Richardson Formula 20 reliabilities are in diagonals.

Factor III in Table 1, Likeability, consists of a clearly homogeneous cluster of items that are inversely related to almost all other items. Wiggins and Winder (8) did not report loadings for these items and may not have included them in their factor analysis.

Product-moment correlations among the four factors derived in the present study and Generalized Kuder-Richardson Formula 20 reliabilities (7) are noted in Table 2. A factor score was calculated for a given subject by adding his raw scores from the variables with factor loadings of .50 and higher. To get the Likeability factor score for a subject, for example, one would add his percentage nomination scores for items 16, 21, 25, 31, 45, 56, 61, and 62. As was true for the Wiggins and Winder data, the Likeability factor is inversely related to the remaining factors. The factor reliabilities range from .86 to .98.

Table 3 depicts the product-moment correlations among the original five PNI scales for the Wiggins and Winder (8) sample and for the current subjects. Reliabilities (7) for the present sample ranged from .82 to .96. The high relationship among the scales, especially between Aggression and Dependency and between Withdrawal and Depression, suggests the need for an item factor analysis. The 58-item factor analysis noted above in Table 1 points to the clustering of items that can be utilized to avoid the overlap and *assumed a priori* independence of variables implied in the five separate scales.

TABLE 3
CORRELATIONS AMONG ORIGINAL WIGGINS AND WINDER SCALES*

PNI scale	Aggression	Dependency	Withdrawal	Depression	Likeability
Aggression (Items 1-12)	.873	.743	.294	.476	—.234
Dependency (Items 13-24)	.819	.926	.467	.620	—.327
Withdrawal (Items 25-36)	.221	.348	.956	.620	—.539
Depression (Items 37-48)	.527	.556	.749	.824	—.412
Likeability (Items 49-56)	—.163	—.077	—.344	—.408	.953

* Upper right half of matrix gives Wiggins and Winder (8) data for 710 boys. Lower left half of matrix has correlations for 113 boys in present study. Diagonals include Generalized Kuder-Richardson Formula 20 reliabilities for 113 boys in present research.

D. SUMMARY

The PNI, a sociometric instrument devised by Wiggins and Winder (8) to measure adjustment in elementary school boys, was evaluated by means of a factor analysis of its 58 items. The item analysis in the present investigation was based on continuously scored items, while the Wiggins and Winder item factor analysis was derived from dichotomously scored items. In general, the results found for the present sample were similar to the Wiggins and Winder findings, although a more clear-cut Aggression-Dependency Factor was indicated in the current research. The high reliabilities, consistency of factor structures, and coverage of important personality characteristics support the potential promise of the PNI as a research tool.

REFERENCES

1. ANASTASI, A. Psychological Testing. New York: Macmillan, 1961.
2. BANDURA, A. Relationship of family patterns to child behavior disorders. Progress Report, United States Public Health Research Grant M-1734, Stanford University, Stanford, California, 1960.
3. HARMAN, H. H. Modern Factor Analysis. Chicago, Ill.: Univ. Chicago Press, 1960.
4. KAISER, H. F. The varimax criteria for analytic rotation in factor analysis. *Psychometrika,* 1958, **23**, 187-200.
5. ———. Comments on communalities and the number of factors. Paper read at conference on "The communality problem in factor analysis," St. Louis, Missouri, May, 1960.
6. SIEGELMAN, M. Loving and punishing parental behavior and introversion tendencies in sons. *Child Devel.,* in press.
7. TRYON, R. C. Reliability and behavior domain validity: Reformulation and historical critique. *Psychol. Bull.,* 1957, **54**, 229-249.
8. WIGGINS, J. S., & WINDER, C. L. The Peer Nomination Inventory: An empirically derived sociometric measure of adjustment in preadolescent boys. *Psychol. Rep.,* 1961, **9**, 643-677.
9. WINDER, C. L., & WIGGINS, J. S. Social reputation and social behavior: A further validation of the Peer Nomination Inventory. *J. Abn. & Soc. Psychol.,* 1964, **68**, 681-684.

The School of Education
The City College of The City
University of New York
New York, New York 10031

[illegible]

[illegible]

1. [illegible] Psychological Testing. New York: Macmillan, 1961.

2. [illegible] patterns in child behavior disorders. [illegible]

[illegible]

9. Wiggins, J. S., & Winder, C. L. [illegible] Social [illegible] and social behavior: A [illegible] study of the Peer Nomination Inventory. [illegible], 1966, [illegible]

Institute of Education
The City College of The City
University of New York
New York, New York 10031

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 151-158.

METHODOLOGY OF RELATIVE SUGAR PREFERENCES IN LABORATORY RATS AND DEER MICE*[1]

Psychology Department, Valparaiso University

Mahlon W. Wagner and J. Timothy Rowntree[2]

A. Introduction

In studying food habits and the origin, development, and modification of food preferences, attention must be given to the methods whereby the preference data are gathered. Stellar and McCleary (6) have compared three different methods of testing food preferences: (*a*) the single-stimulus method; (*b*) the two-stimulus method, one being water; and (*c*) the two-stimulus method with two different concentrations. These methods were used to measure separately the rat's preference for glucose and for sodium chloride. The preference for sodium chloride was unaffected by the method used, being maximally preferred at 0.8 per cent concentration. The maximal preference for glucose was 5 per cent by the single stimulus method, 10 per cent when paired with water, and 20 per cent when paired with another sugar solution. Jacobs (2) compared the latter two methods using glucose concentrations of 4.5, 13.5, and 36 per cent. Using glucose and water as choices, Jacobs found 4.5 per cent and 13.5 per cent equally preferable and 36 per cent less preferred in a six-hour test. During the 18-hour period immediately following the short-term test, "a relative preference for the 4.5 per cent solution in relation to the other two solutions appeared with an inverse relationship between concentration and intake" (2, p. 1045). Using two glucose solutions as the choice, initial preference was for the sweeter, but as the rats continued to drink the relation shifted such that the 13.5 per cent solution was the more preferred.

From these studies it is seen that the choices available determine the relative preference and, further, it seems as if duration of drinking is a factor affecting relative choice. With respect to the effects of satiation, to which Jacobs ascribes the preference shift, Wagner (7, 8) did not observe any shift of this type with prolonged exposure to 12.5 per cent and 25 per cent glucose solutions.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on August 15, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by funds from an NSF Institutional grant to Valparaiso University and NIH Grant MH 10185 to Mahlon W. Wagner.

[2] Now at Boeing Company, Seattle, Washington.

If more than two different concentrations are to be compared, an additional problem arises: should all *S*s be exposed to all concentration pairs or should separate groups be used for each experimental treatment? Favoring the latter method, Young (10) has pointed out that a preference develops only through experience and practice and, once formed, tends to persist even in the presence of conflicting metabolic need. Harriman (1) also used prior experience to establish a strong sugar-over-salt preference, which persisted after the rats were adrenalectomized to the extent that four of 20 rats died and the others lost weight. Wagner (8) found that 14 days of experience facilitated the development of a clear preference for a 25 per cent glucose solution over a 12.5 per cent solution. Jacobs (2) does not make clear which of the above two methods was used to obtain relative preferences for the three glucose solutions. However, Wagner, Green, and Manley (9) presented all possible combinations of five glucose solutions in a paired comparisons procedure to examine the sugar preference of Squirrel monkeys and found no changes in relative preference on a second replication of the pairings.

The purpose of the present experiment was to compare these two methods of evaluating relative sugar preference.

B. METHOD

1. *Subjects*

Two strains of male laboratory rats, Long-Evans hooded and Wistar albino, and two species of White-footed Deer mice of the family Cricetidae, *Peromyscus leucopus* and *Peromyscus maniculatus bairdii,* generously supplied by John A. King from the Animal Behavior Laboratory at Michigan State University, served as subjects.

2. *Apparatus*

At all times *S*s were housed individually in wire mesh cages measuring 14 × seven × six inches. Three concentrations of glucose solution, 10, 25, and 37.5 per cent, were prepared on a weight-by-volume basis following the recommendations of Pfaffmann, Young, Dethier, Richter, and Stellar (4). All solutions were presented in inverted graduated test tubes fitted with glass drinking spouts of two mm inside diameter.

3. *Procedure*

All rat strains and mice species received identical treatment and therefore only the general experimental design is presented. One group of five randomly assigned *S*s received all three pairs, 10-25, 10-37.5, and 25-37.5 per cent, on a randomly cycled schedule repeating each pair five times over 15 days. Three groups of five *S*s each received the same pair of solutions each day for 15 days (noncycling). Glucose solutions were presented to the mice for two hours per day between 9 AM and 1 PM, with intake recorded every 30 minutes to the nearest one-fourth cc. All rat *S*s were given 30 minutes exposure to the two solutions per day. The differences in exposure time were necessary because the much smaller mice ingested proportionately less fluid per unit time. Standard laboratory chow was always available and water was present, except during preference testing. Solution positions were randomly altered to prevent development of spatial position preference.

C. RESULTS

1. *Method Differences for Mice*

Relative sugar preference (sweeter solution minus less sweet solution) was determined for all three pairs of solutions and for both species of White-footed Deer mice. In order to compare the two methods, the data for noncycled *S*s are presented as three-day means, giving five points, and plotted with the five pairings for that same concentration pair presented to the cycled *S*s. There were no species differences at any of the three concentration pairs and, therefore, the data have been pooled for species. All analyses of variance reported below used actual intake rather than relative intake as data.

Relative preference for the 10 per cent *vs.* 25 per cent glucose solutions is presented in Figure 1a. An analysis of variance indicates that the two methods yield significant intake differences ($p < .02$), methods yield different intake over trials ($p < .002$), and the following effects were also significant: 25 per cent preferred to 10 per cent ($p < .002$), trials ($p < .002$), trial × concentration ($p < .002$), and method × trial × concentration ($p < .05$).

Relative preference for the 10 per cent *vs.* 37.5 per cent glucose solutions is presented in Figure 1b. For the mice a definite preference for the sweeter solution is evident ($p < .002$). The main effect of method was not significant, but the interaction of method × trial was $p < .01$. Intake over trials was also significant ($p < .002$).

Data concerning the two sweetest solutions, 25 and 37.5 per cent, are presented in Figure 1c. Cycling produced increased intake of the sweeter solution (method, $p < .05$) and more pronounced intake differences over trials (method × trial, $p < .01$). Again the sweeter solution was definitely preferred (concentration, $p < .01$), intake changed over trials (trials, $p < .002$), and degree of preference changed over trials (trial × concentration, $p <$

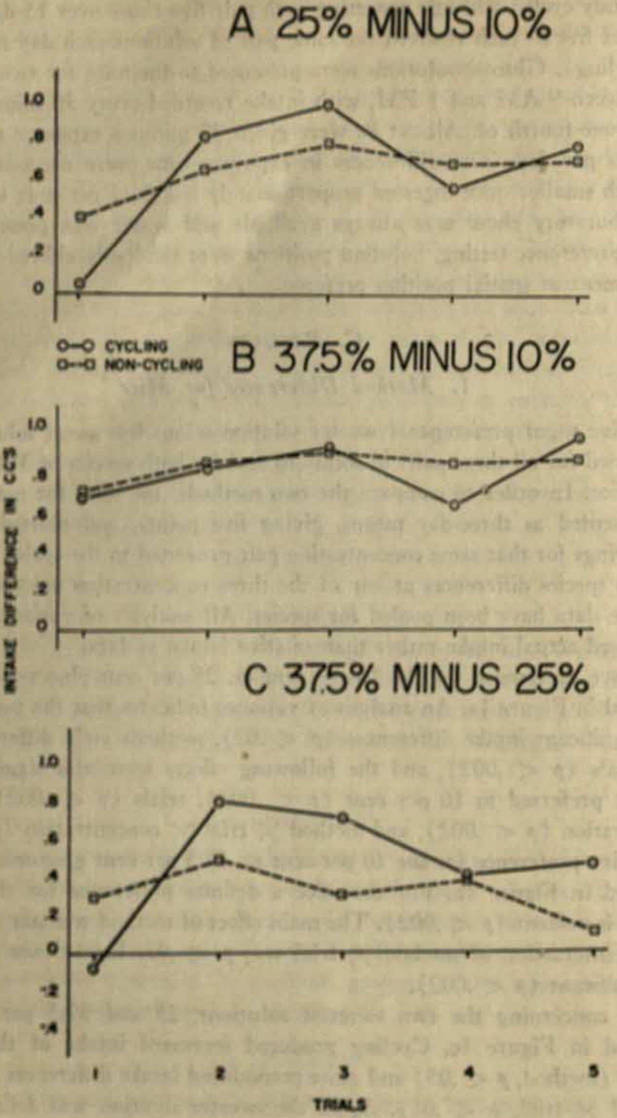

FIGURE 1

RELATIVE SUGAR PREFERENCE OF DEER MICE WITH TWO METHODS OF PRESENTATION OF SUGARS

.05). There were also changes in preference over trials due to method approaching significance (method × trial × concentration, $p = .07$).

2. *Method Differences for Rats*

Between the two strains of rats there were no significant differences, and the intakes for the Long-Evans and Wistar *Ss* were pooled. The preferences for the glucose solutions are shown in Figures 2a, b, and c, for the 10-25 per cent, 10-37.5 per cent, and 25-37.5 per cent pairs, respectively. The rats significantly preferred the 25 per cent to the 10 per cent (concentration, $p < .01$) and this preference changed over trials (trials × concentration, $p < .002$). The category of trials alone was also significant ($p < .002$). The rats also preferred the 37.5 per cent solution to the 10 per cent, but to a lesser degree (concentration, $p < .05$), and again preference changed over trials ($p < .05$).

There were no significant preferences by the rats for the sweeter of the 25-37.5 per cent pair, although total fluid intake was no different from the other two solutions. The only significant effect was a change in intake over trials ($p < .002$).

D. DISCUSSION

The two experimental designs, of noncycling (separate groups of animals for each pair of solutions) and cycling (subject by treatment, where every animal was exposed to all three pairs of solutions on a randomly rotated schedule), were evaluated in order to determine if one method produced larger or more clear-cut differences in preference, if these differences would be uniform for all pairs of solutions or particular to only certain pairs, and if the species might exhibit different preferences with the two methods.

Of particular interest is the development of preference for the pair of sweetest solutions (25 per cent *vs.* 37.5 per cent). Classical psychophysical law would predict that this pair of solutions, more than any other, would be the most confusing and give the most inconsistent data for the animals, since in sensory terms these solutions would be the most indiscriminable in terms of relative sweetness, producing the lowest ratio of $\Delta I/I$ (for 10 *vs.* 25 per cent, $\Delta I/I = 1.5$; for 10 *vs.* 37.5 per cent, $\Delta I/I = 2.7$; and for 25 *vs.* 37.5 per cent, $\Delta I/I = 0.5$). For both strains of rats the 25-37.5 per cent pair were not discriminated, and for both species of Deer mice the preference discrimination was less significant than for the other two pairs. Further, for the Deer mice the two methods produced different intake, different intake over trials, and changing preference over trials. The cycling method provided the greater preferences and the greater preference changes over trials.

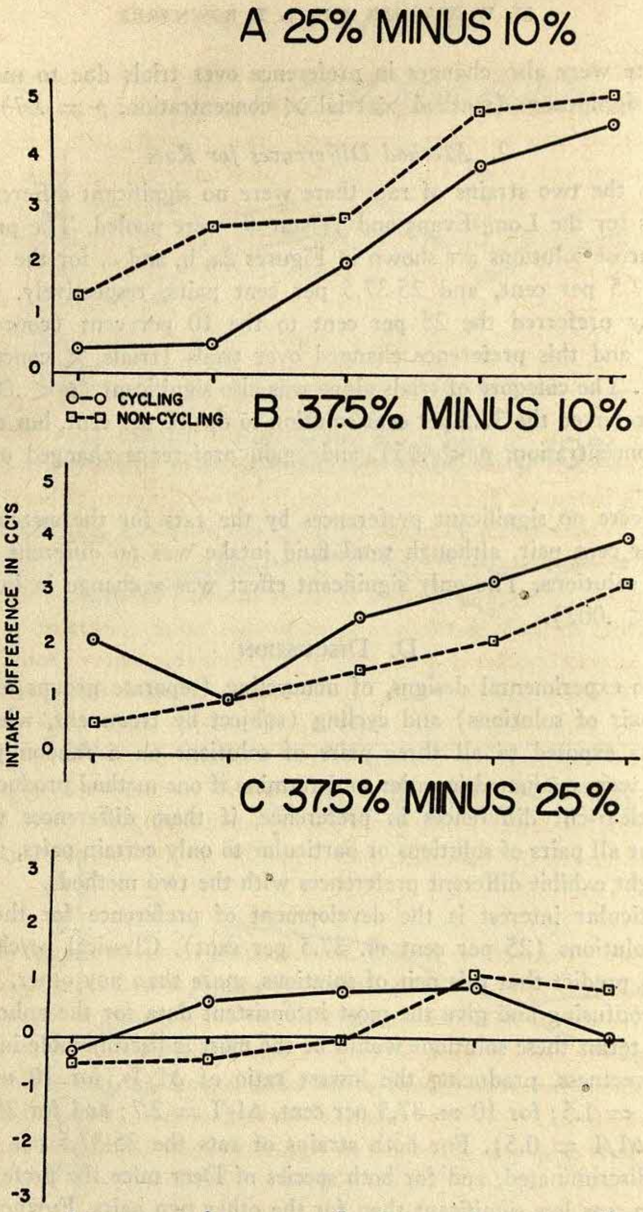

FIGURE 2
RELATIVE SUGAR PREFERENCE OF LABORATORY RATS WITH TWO METHODS OF PRESENTATION OF SUGARS

Of the two remaining pairs of solutions, only with the 10-25 per cent pair was a method effect significant. Again, as with the 25-37.5 per cent pair, the cycling method was superior, producing greater intake differences, greater differences over trials, and a more marked preference change with experience.

The question of why mice show intake and preference differences as a result of being exposed to the two presentation methods, while the laboratory rats show no effect of method, cannot be resolved by this study. It is true that the mice are much closer in terms of time and genetics to the wild state, whereas the rats have been laboratory reared and genetically inbred for countless generations. Richter (5) dramatically showed that domestication in rats can indeed reduce their reactivity to stimuli and the efficiency of their discriminatory abilities. Also the rats here were under no stress or other imposed experimental conditions that might serve to enhance any latent constitutional differences. McEwen (3) reports strain differences between Long-Evans and Wistar strain rats, but by using electric shock and also examining for sex differences.

In summary, the cycling method is either superior to, or no different from, the noncycling method in producing larger and more consistent preferences. Cycling is especially superior with the less discriminable pairs of solutions. Consequently, it seems that the cycling procedure (a paired-comparisons method) should be recommended in future testing of relative sugar preference. Additional factors favoring the cycling method include the use of the animal as his own control, the provision of intraanimal comparisons of preference, and the need for fewer animals with its related savings in time. Individual differences can also be studied in more detail by the cycling method.

E. SUMMARY

This study examined preference development using two methods of presentation of sugar solutions. Two species of White-footed Deer mice and two strains of laboratory rats were used. Pairs of sugar solutions were presented either in cycling fashion, each *S* being exposed to all possible pairs, or in noncycling fashion, different groups of *S*s being exposed to the different pairs. Cycling was found to be the superior method, both in terms of faster and clearer preference development and in terms of experimental and analysis time and effort.

REFERENCES

1. HARRIMAN, A. The effects of a pre-operative preference for sugar over salt upon compensatory salt selection by adrenalectomized rats. *J. Nutrition,* 1955, **57**, 271-276.
2. JACOBS, H. L. Some physical, metabolic, and sensory components in the appetite for glucose. *Amer. J. Physiol.,* 1962, **203**, 1043-1054.
3. MCEWEN, B. Constitution, experience and alcohol intake in the rat. Paper presented at Eastern Psychological Association Meeting, Atlantic City, New Jersey, 1965.
4. PFAFFMANN, C., YOUNG, P. T., DETHIER, V. G., RICHTER, C. P., & STELLAR, E. The preparation of solutions for research in chemoreception and food acceptance. *J. Comp. & Physiol. Psychol.,* 1954, **47**, 93-96.
5. RICHTER, C. P. Rats, man and the welfare state. *Amer. Psychol.,* 1959, **14**, 18-28.
6. STELLAR, E., & MCCLEARY, R. A. Food preference as a function of the method of measurement. *Amer. Psychol.,* 1952, **7**, 256 (Abstract).
7. WAGNER, M. W. Satiation effects on rate of intake and preference of glucose solutions. *J. Comp. & Physiol. Psychol.,* 1965, **59**, 115-117.
8. ————. The effects of age, weight and experience on relative sugar preference in the albino rat. *Psychon. Sci.,* 1965, **2**, 243-244.
9. WAGNER, M. W., GREEN, K. F., & MANLEY, M. B. Paired comparison method for measurement of sugar preference in Squirrel monkeys. *Science,* 1965, **148**, 1473-1474.
10. YOUNG, P. T. The role of affective processes in learning and motivation. *Psychol. Rev.,* 1959, **66**, 104-125.

Psychology Department
Valparaiso University
Valparaiso, Indiana 46383

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 159-166.

FACTORS AND FACTOR ANALYSES*[1]

Department of Psychology, University of California, Los Angeles

J. A. GENGERELLI

A. FACTOR ANALYSIS OF A CORRELATION MATRIX GENERATED BY CONSTELLATIONS OF KNOWN UNIT DETERMINERS

The basic motives for doing a factor analysis are (*a*) to determine the number of and to identify the separate components involved in a set of psychological tests, however large; and (*b*) to estimate the value of each of the components in each member of the set. Current factor analytic procedures achieve these aims, except that the components are not identified. Given a set of n tests, the factor analysis tells us that there exists a set of orthogonal components (factors), c in number ($c \ll n$), which are capable of generating the correlation matrix; however, we are not told the nature of these dimensions nor how they could be realized in practice. An effort is made to meet the requirement by rotating the axes, but this maneuver typically leads to results that are only suggestive at best and very often merely give ingenious verbal tags.

This outcome is to be expected. If we look at the n tests in the usual fashion as n vectors arranged in a c-space in such wise as to reflect the correlations obtaining between them, the c axes (factors) are reference vectors in terms of which the content of each of the n tests is described. But the relationship between the test vectors and the reference vectors is symmetrical. Thus, if a certain component of a given test vector is described in terms of a given reference vector, it is also the case that the constitution of that reference vector is describable *in part* by the constitution of the test vector. This situation will be true for each of the n test vectors vis-à-vis that reference vector. The mathematical machinery, in striving to account for the correlation matrix with the minimal number of reference vectors, produces vectors of mixed composition. Hence, instead of being "pure" and "homogeneous," the reference vector may be a *mélange* of a number of distinct psychological functions. It is as if we constructed a test constituted (to take a farfetched illustration)

---

* Received in the Editorial Office, Provincetown, Massachusetts, on August 22, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Supported by a grant from the University of California, Los Angeles, Faculty Research Committee.

of reaction time, auditory memory span, paired-associates learning, and critical flicker-frequency; summed the scores in an acceptable fashion; and had a single vector represent it as a "factor." It would be difficult, indeed, to name such a vector.

These difficulties in multiple factor analysis were, in fact, pointed out from the very beginning [e.g., Thurstone (9, p. 179)]; nonetheless, many investigators seem to ignore them, at least by implication, in their work.

By way of illustration, a principal axes factor analysis has been applied to a set of synthetic data in which the components (factors) are (*a*) known and (*b*) homogeneous.[2] In Table 1 is given the "microstructure" of a set of 20 test variables, of which six mutually correlate zero and the other 14 are constituted of two or more of these six. Each of the 20 tests is constituted of varying numbers of unit determiners, the range being 4-24. Utilizing the well-known common elements formula, the correlation matrix given in Table 2 was obtained. Application of the minimum residual factor analytic technique [Comrey (1, 2)] yielded the factor matrix shown in Table 3. The matrix

TABLE 1

MICROSTRUCTURE OF 20 TESTS

Test	I	II	III	IV	V	VI	Unit determiners
(1)	XXXX	XXXX	——	——	XXXX	——	12
(2)	——	——	XXXX	XXXX	——	XXXX	12
(3)	XXXX	XXXX	XXXX	XXXX	——	——	16
(4)	XXXX	——	——	——	——	——	4
(5)	——	——	——	XXXX	XXXX	——	8
(6)	——	XXXX	XXXX	——	——	——	8
(7)	——	XXXX	——	——	——	——	4
(8)	——	——	XXXX	XXXX	XXXX	XXXX	16
(9)	——	XXXX	XXXX	XXXX	XXXX	XXXX	20
(10)	——	——	XXXX	——	——	——	4
(11)	XXXX	——	XXXX	XXXX	XXXX	XXXX	20
(12)	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	24
(13)	——	——	——	XXXX	——	——	4
(14)	——	——	XXXX	XXXX	——	——	8
(15)	XXXX	——	——	XXXX	——	XXXX	12
(16)	——	——	——	——	XXXX	——	4
(17)	——	——	——	XXXX	XXXX	——	8
(18)	XXXX	——	——	——	——	XXXX	8
(19)	——	——	——	——	——	XXXX	4
(20)	——	——	XXXX	XXXX	XXXX	——	12

---

[2] This work was undertaken and finished before Overall's interesting paper (7) came to the writer's attention.

TABLE 2
INTERCORRELATIONS AMONG THE 20 TESTS

Test	1	2	3	4	5	6	7	8	9	10	11
1	1.000	0	.577	.577	.408	.408	.577	.289	.516	0	.516
2	0	1.000	.577	0	.408	.408	0	.866	.774	.577	.774
3	.577	.577	1.000	.500	.354	.707	.500	.500	.667	.500	.667
4	.577	0	.500	1.000	0	0	0	0	0	0	.447
5	.408	.408	.354	0	1.000	0	0	.707	.632	0	.632
6	.408	.408	.707	0	0	1.000	.707	.354	.632	.707	.316
7	.577	0	.500	0	0	.707	1.000	0	.447	0	0
8	.289	.866	.500	0	.707	.354	0	1.000	.894	.500	.894
9	.516	.774	.667	0	.632	.632	.447	.894	1.000	.447	.800
10	0	.577	.500	0	0	.707	0	.500	.447	1.000	.447
11	.516	.774	.667	.447	.632	.316	0	.894	.800	.447	1.000
12	.707	.707	.816	.408	.577	.577	.408	.816	.913	.408	.913
13	0	.577	.500	0	.707	0	0	.500	.447	0	.447
14	0	.816	.707	0	.500	.500	0	.707	.632	.707	.632
15	.333	.667	.577	.577	.408	0	0	.577	.516	0	.774
16	.577	0	0	0	.707	0	0	.500	.447	0	.447
17	.408	.408	.354	0	1.000	0	0	.707	.632	0	.632
18	.408	.408	.354	.707	0	0	0	.354	.316	0	.632
19	0	.577	0	0	0	0	0	.500	.447	0	.447
20	.333	.667	.577	0	.816	.408	0	.866	.774	.577	.774

TABLE 2 (*continued*)

Test	12	13	14	15	16	17	18	19	20	Σ
1	.707	0	0	.333	.577	.408	.408	0	.333	7.634
2	.707	.577	.816	.667	0	.408	.408	.577	.667	10.211
3	.816	.500	.707	.577	0	.354	.354	0	.577	10.434
4	.408	0	0	.577	0	0	.707	0	0	4.216
5	.577	.707	.500	.408	.707	1.000	0	0	.816	8.856
6	.577	0	.500	0	0	0	0	0	.408	6.724
7	.408	0	0	0	0	0	0	0	0	3.639
8	.816	.500	.707	.577	.500	.707	.354	.500	.866	11.531
9	.913	.447	.632	.516	.447	.632	.316	.447	.774	11.933
10	.408	0	.707	0	0	0	0	0	.577	5.870
11	.913	.447	.632	.774	.447	.632	.632	.447	.774	12.191
12	1.000	.408	.577	.707	.408	.577	.577	.408	.707	12.619
13	.408	1.000	.707	.577	0	.707	0	0	.577	7.154
14	.577	.707	1.000	.408	0	.500	0	0	.816	9.209
15	.707	.577	.408	1.000	0	.408	.816	.577	.333	9.155
16	.408	0	0	0	1.000	.707	0	0	.577	5.370
17	.577	.707	.500	.408	.707	1.000	0	0	.816	8.856
18	.577	0	0	.816	0	0	1.000	.707	0	6.279
19	.408	0	0	.577	0	0	.707	1.000	0	4.663
20	.707	.577	.816	.333	.577	.816	0	0	1.000	10.618

$r_t = 167.162$

TABLE 3
MINIMUM RESIDUAL FACTOR MATRIX

Test	1	2	3	4	5	6	7	h^2
1	0.524	0.335	—0.017	0.781	—0.009	—0.056	0.004	1.000
2	0.818	0.019	0.104	—0.548	0.102	0.092	0.032	1.000
3	0.780	0.231	0.390	0.123	—0.412	0.029	0.035	1.001
4	0.266	0.682	—0.191	0.231	—0.470	—0.392	—0.009	1.000
5	0.724	—0.481	—0.416	0.240	—0.077	0.088	0.012	1.000
6	0.490	0.088	0.840	0.178	0.116	0.044	0.008	1.001
7	0.221	0.233	0.538	0.543	0.047	0.558	—0.000	1.000
8	0.920	—0.150	—0.097	—0.185	0.293	—0.052	0.010	1.001
9	0.920	—0.031	0.153	0.077	0.283	0.203	—0.044	1.000
10	0.472	—0.108	0.650	—0.291	0.116	—0.495	—0.004	1.000
11	0.940	0.169	—0.173	—0.062	0.053	—0.222	—0.051	1.000
12	0.949	0.250	0.062	0.164	0.066	0.025	—0.014	1.000
13	0.602	—0.346	—0.214	—0.239	—0.515	0.387	—0.019	1.000
14	0.759	—0.320	0.309	—0.375	—0.283	—0.076	—0.014	1.000
15	0.699	0.474	—0.381	—0.247	—0.236	0.152	0.007	0.999
16	0.421	—0.336	—0.375	0.578	0.407	—0.263	0.009	1.000
17	0.724	—0.481	—0.416	0.240	—0.077	0.088	0.012	1.000
18	0.432	0.826	—0.316	—0.133	0.075	—0.088	0.020	1.001
19	0.344	0.486	—0.256	—0.419	0.576	0.267	0.007	1.000
20	0.862	—0.456	0.036	0.028	0.004	—0.214	0.027	1.000

TABLE 4
ROTATED FACTOR MATRIX (VARIMAX CRITERION)

Test	1	2	3	4	5	6	7	h^2
1	0.441	—0.004	—0.045	0.598	0.286	0.604	—0.006	1.000
2	0.278	0.616	0.609	0.015	—0.413	0.027	0.030	1.000
3	0.216	0.030	0.512	0.506	—0.397	0.525	0.042	1.001
4	—0.048	—0.005	—0.004	0.999	—0.010	0.007	—0.004	1.000
5	0.973	0.037	0.038	0.046	—0.219	0.028	0.005	1.000
6	—0.038	0.009	0.672	—0.004	0.040	0.739	0.006	1.001
7	—0.033	—0.003	—0.046	—0.009	—0.027	0.998	—0.001	1.000
8	0.662	0.534	0.515	0.036	—0.091	0.045	—0.002	1.001
9	0.577	0.475	0.438	0.028	—0.095	0.486	—0.055	1.000
10	—0.022	0.015	0.995	0.004	0.083	0.047	—0.007	1.000
11	0.570	0.474	0.457	0.478	—0.086	0.043	—0.059	1.000
12	0.508	0.432	0.400	0.434	—0.090	0.446	—0.022	1.000
13	0.532	0.052	0.081	0.018	—0.841	—0.002	—0.007	1.000
14	0.360	0.048	0.761	0.016	—0.536	0.032	—0.008	1.000
15	0.263	0.603	0.033	0.589	—0.467	0.004	0.009	0.999
16	0.844	0.000	—0.028	0.045	0.532	0.041	—0.013	1.000
17	0.973	0.037	0.038	0.046	—0.219	0.028	0.005	1.000
18	—0.054	0.703	—0.017	0.710	0.023	0.007	0.015	1.001
19	—0.028	0.998	—0.020	0.004	0.042	0.002	—0.006	1.000
20	0.781	0.039	0.605	0.039	—0.130	0.050	0.020	1.000

was then rotated analytically, using the normal varimax criterion [Kaiser (6)], to give the rotated factor matrix presented in Table 4.[3] The "true" factor matrix is exhibited in Table 5.

A comparison of Tables 4 and 5 is, to say the least, disenchanting. If we ignore the seventh column of factor loadings in Table 4, in which the entries are uniformly small, the analysis yielded the requisite six factors; this is on the credit side. Further, the variance of tests 4, 7, 10, and 19 is concentrated (for all practical purposes) in but one factor; comparison with Tables 1 and 5 shows this to be indeed the case. The analysis goes awry, however, in Tests 13 and 16. These are single factor tests, but the rotated factor matrix shows them to have substantial loadings in two factors. *Per contra,* Tests 17 and 5, which Tables 1 and 5 show as two-factor tests, are given a very heavy loading on only one factor. Test 12, which "in reality" has equal loadings on all six factors, is given loadings on only five; Test 11, with equal loadings on five factors, is given loadings on only four.

TABLE 5

"TRUE" FACTOR COMPOSITION OF THE 20 TESTS

Test	I (4)	II (7)	III (10)	IV (13)	V (16)	VI (19)
1	.577	.577	—	—	.577	—
2	—	—	.577	.577	—	.577
3	.500	.500	.500	.500	—	—
4	1.000	—	—	—	—	—
5	—	—	—	.707	.707	—
6	—	.707	.707	—	—	—
7	—	1.000	—	—	—	—
8	—	—	.500	.500	.500	.500
9	—	.447	.447	.447	.447	.447
10	—	—	1.000	—	—	—
11	.447	—	.447	.447	.447	.447
12	.408	.408	.408	.408	.408	.408
13	—	—	—	1.000	—	—
14	—	—	.707	.707	—	—
15	.577	—	—	.577	—	.577
16	—	—	—	—	1.000	—
17	—	—	—	.707	.707	—
18	.707	—	—	—	—	.707
19	—	—	—	—	—	1.000
20	—	—	.577	.577	.577	—

Note: The number in parenthesis at the top of each column identifies the factor.

---

[3] The writer thanks Professor A. L. Comrey for his kindness in carrying out the calculations on the IBM 7094 operated by the University of California, Los Angeles, Computing Facility.

The results obtained with these synthetic data bring out the intrinsic psychological ambiguities involved in the vectorial manipulations underlying factor analysis techniques. The findings reported here are in agreement with those given in Overall's report (7), and the comments made in that paper apply here. Factor analysis becomes a useful technique only when it goes hand-in-hand with empirical test construction and modifications dictated by psychological theory and antecedent factor analytic findings. Thus an original set of *n* tests, assembled under the guidance of a well-thought-out hypothesis, should be reconstituted on the basis of successive factor analyses until a final set *n′* is obtained whose factors have operational meaning [Guilford (5)].

B. Analysis of the Same Matrix in Terms of Its Component Parts

It is always possible to describe the composition of a set of *n* tests in terms of specific tests contained in the set (3). However, in the context of psychometric analysis, we require that these tests be mutually orthogonal. Since this condition is practically never met, we can compromise by describing the set *n* in terms of a subset whose members deviate from mutual orthogonality by some arbitrary value, $r_{xy} \overline{<} k$. When *n* is large, the isolation of a subset of tests whose members do not intercorrelate more than *k* presents difficulties, but a routine method for finding such subsets has been described (4). When the subset has been isolated, a regression equation is constructed with the tests in the subset as the predictor variables and the sum of the scores of the total set *n* as the predicted variable.

The correlations between each of the predictor variables and the sum of the *n* tests necessary for computing the required regression coefficients is given by

$$r_{js} = \frac{\sum_{x=1}^{n} r_{jx}}{(r_t)^{\frac{1}{2}}},$$

where r_t is the sum of the entries in the full correlation matrix (1 in the diagonal cells) and the numerator is the sum of the entries in the *j*th column (see Table 2). Computation of the β values is accomplished by utilizing the formula

$$P^{-1}\,C = B$$

programmed for a computer [Ray (8, pp. 166-9)]. Here *P* is the full correlation matrix (with 1 in the diagonal) of the predictor variables, C is the column matrix containing the correlations between each of the predictor variables and the sum, and B is the column matrix of the desired regression coefficients.

If the regression coefficients thus obtained yield a satisfactory multiple correlation, we may consider that the subset of tests describes to an acceptable degree of approximation the entire set; otherwise, we must employ a criterion of intercorrelation larger than k, say l, and repeat the process.

Once it is shown that a specific subset of empirical tests is capable of predicting the variance of the entire set (within the limits of acceptability imposed), it is possible to derive a factor matrix in which the factors are the tests in the subset. This is accomplished by constructing n multiple regression equations, one for each of the n tests. In these equations the criterion variable is, in turn, the test to be analyzed and the predictor variables are, in each instance, the tests comprising the factors. The β coefficients obtained in each equation represent the loadings of the factors involved in the test in question.

The matrix procedure for computing the β's is the same as before; the only change is in the contents of column matrix C, which will vary from test to test.

When the above procedure is applied to the present synthetic data, with $k = 0$, it is found (of course) that there is a subset of six tests—namely, 4, 7, 10, 13, 16, and 19—which predicts the variance of the sum of the 20 tests completely. The multiple regression coefficients are $\beta_4 = .326$, $\beta_7 = .281$, $\beta_{10} = .454$, $\beta_{13} = .553$, $\beta_{16} = .415$, and $\beta_{19} = .361$. Since the intercorrelations between members of the subset are all zero, the factor loadings exhibited in Table 5 are obtained by inspecting Table 2: the loading that Factor j has on Test x is equal to r_{jx}.

The method that has been outlined and illustrated offers attractive possibilities. What is required for the factorial analysis of any set of n psychological tests is the availability of another set of tests that intercorrelate .10 or less with one another. By incorporating this latter set into the original battery of n tests and applying the above procedure, we determine to what extent the battery is described by the (quasi) orthogonally interrelated tests; if this is satisfactory, we proceed to determine the factor loadings. This implies that systematic efforts should be made to cull from the literature, and to construct, psychological tests whose mutual intercorrelations are negligible (3). The devising of tests of psychological functions that intercorrelate negligibly presents an important challenge to psychology. It makes operational realities of "factors" that otherwise must remain abstract conceptual entities.

REFERENCES

1. COMREY, A. L. The minimum residual method of factor analysis. *Psychol. Rep.*, 1962, **11**, 15-18.
2. COMREY, A. L., & AHUMADA, A. An improved procedure and program for minimum residual factor analysis. *Psychol. Rep.*, 1964, **15**, 91-96.
3. GENGERELLI, J. A. A method of analysis in which the factors are empirical tests. *J. of Psychol.*, 1952, **33**, 159-174.
4. ———. A method for detecting subgroups in a population and specifying their membership. *J. of Psychol.*, 1963, **55**, 457-468.
5. GUILFORD, J. P. The structure of the intellect. *Psychol. Bull.*, 1956, **53**, 267-293.
6. KAISER, H. F. The varimax criterion of analytic rotation in factor analysis. *Psychometrika,* 1958, **23**, 187-200.
7. OVERALL, J. E. Note on the scientific status of factors. *Psychol. Bull.*, 1964, **61**, 270-276.
8. RAY, W. S. Statistics in Psychological Research. New York: Macmillan, 1962.
9. THURSTONE, L. L. Multiple Factor Analysis (2nd ed.). Chicago: Univ. of Chicago Press, 1947.

Department of Psychology
University of California, Los Angeles
Los Angeles, California 90024

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 167-191.

JUST HOW FAR CAN SOCIAL CHANGE CHANGE PERSONALITY?*[1]

Louisiana State University

RALPH MASON DREGER

A. INTRODUCTION

Social psychology has been interested for many years in changes of attitudes under changing social conditions, and has developed a number of theories to account for the alteration or lack of alteration of attitudes under varying conditions (23). The more general problem—of change under the impact of changing social circumstances in personality, of which attitudes are only a part—has but recently been given explicit attention [e.g., Worchel and Byrne (89)]. It is to this more general problem that this paper addresses itself.

For the sake of those who are not familiar with the eclectic "working model" of the personality that the writer utilizes in both his theory and clinical practice, the following review capsulates the position set forth in the writer's text on personality (25). According to this view, personality consists of the organized functions of the individual. This definition is general enough that it can be adapted to any overall theory of personality. If one goes along with Gordon Allport (1, 2) and regards personality as contained within the skin, he may do so. If one is more sociologically oriented—for example, with Gardner Murphy (58) or Harry Stack Sullivan (84)—a concept of personality as partly interpersonal can utilize the definition. The author's own theory incorporates both "within the skin" concepts and roles; he cannot see how any understanding of individuality is possible if roles are considered as external to the personality, but he recognizes that not all authorities are in agreement on this point [see Allport (2)].

B. THE MODEL OF THE PERSONALITY

The above concepts are encompassed in overlapping sets derived from various personality theories, all of which constructs serve as efforts to order

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 1, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This article is adapted from a lecture delivered at Clark College, May 3, 1966, in the Clark College Winter and Spring Lecture Series 1966, "The Social Sciences and the Development of Human Potentials."

the amazingly complex data of personality. First, use is made of an expanded Freudian scheme of dynamic structure—especially id, ego, and superego—as one classification of functions, and conscious-unconscious as another. Like others [e.g., Kluckhohn, Murray, and Schneider (47)] who have enlarged Freud's original notions, the writer has come to think of the *id* as the dynamic reservoir not only of evil intentions or amoral impulses, but also of intents acceptable to and encouraged by society, such as love, worship, air hunger, food hunger, and all metabolic processes or tissue needs. Presumably the inherited component is a large portion of the id. The drives or primary impulses are initially DNA (deoxyribonucleic acid) derived; we can say they are at the heart of the id as the id is at the heart of the personality.

Ego, which arises as a set of control functions to master and direct id energies and to gain control over stimuli arising outside of the organism, consists of at least five major subfunctions: perception, mastery of motility or of the motor apparatus, "binding" or controlling tension, judging or thinking, and synthesizing or attempting to harmonize the bewildering and often conflicting functions of the personality and aspects of the external world.

Last in Freud's schema, but not in the writer's, is the *superego,* a "precipitate within the ego" as Freud said. Superego is the set of functions developed as controls for the personality relating to the demands of society—both custom and conscience demands (27)—mediated first by the parents or parent-surrogates and later by siblings, peers, teachers, and others who can give or withhold love.

Finally, in this simplified classification of personality functions, there are the *roles* we enact in the world. These are interobject, mostly interpersonal, portions of the personality. If one thinks that roles are not organized functions of the individual and therefore parts of the personality, he may not have observed, as most clinicians have, parents and children in symbiotic relation to one another; he may not have perceived, although he may have seen, husband and wife combinations wherein the death of one causes the other to say, quite accurately, "Part of me died when he died." The roles of parent and child and the roles of truly married husband and wife are *inter*personal, not within the skin. Since the general topic of this paper relates to changes in personality accompanying social changes, the concept of role already commits the writer to the contention that changes in roles, mostly social changes, *are* personality changes, inasmuch as roles are integral portions of the dynamic structure of personality.

Going beyond id, ego, superego, and roles, again with Freud and a host of others, the theory of personality reviewed here cross-classifies personality functions into *conscious* and *unconscious.* As is mostly obvious, consciousness and unconsciousness overlap all other functions. Thinking and judging, affects, and ambitions may be either conscious or unconscious. By definition the hidden source of the personality's energies, the id, is unconscious. But all other functions may be either or both conscious and unconscious—even roles. Again, if one thinks roles are totally conscious, one has not observed a mother's reinforcing her child's dependency by subtle though entirely recognizable cues—mother role behaviors of which she is mostly unaware—or the role behaviors of Negro and white in interactions that they dimly perceive in themselves, if at all. One must also become familiar with the scientific literature on unconscious factors involved in enacting a role. This literature is too extensive to cite [*cf.* Sarbin (73, 74)], but it goes back at least to the theories of G. H. Mead (56) and the experimental studies of Solomon Asch (5) in the social modification of autokinetic behavior.

Some theorists, one must frankly admit, deny the reality of the conscious-unconscious classification of personality functions. As far as can be ascertained, however, no one denies the reality of the distinction between central and peripheral functions, a cross-classification suggested most lucidly by Kurt Lewin (51). All that is needed is to call attention to phenomena with which we are all familiar to clarify what is meant here. The well-known congressional hearings on the Ku Klux Klan saw a procession of witnesses who, in effect, said, "I know, but I ain't sayin'." Central functions are those that we are unwilling or unable to reveal to others; peripheral functions are those that we are able or willing to disclose to others.

Again, another overlapping classification of personality functions divides them into *self* and *not-self* or more precisely, in the writer's view, into *selves* and *not-selves.* There are those parts of our personalities, including our roles, that we accept as our own, and those that we reject. The former we subsume under a composite rubric, the self. Although it seems to go contrary to common sense and to the notions of many theorists to maintain that we have selves rather than a single self, the burden of proof rests on those who assert that all the disparate and contradictory functions we claim as "I," "me," or "mine" constitute a single, monolithic self. Roughly, though, we may say that personality can be divided into self and not-self.

One last category of functions completes the present image of the personality. Functions of the personality can be cross-classified into *traits,* sometimes called sets or dispositions or sentiments, or in the Allportian sense just plain traits of personality. These may be conscious or unconscious, or more likely both. They may be central or peripheral, or likewise more likely both.

And they may share components of id, ego, superego, and roles. If one takes, for example, the trait of dominance-submissiveness, either factor-analytically derived or derived from a single scale test, it is not impossible by any means to show that it has drive (or id) qualities, it has some measure of rational (ego) control processes, some superego aspects, and some role-taking properties, that it seems to possess conscious and unconscious characteristics, that it almost certainly is both central and peripheral, and that some features of this trait are acceptable as "mine," while others may be rejected as "not-mine." Traits like dominance-submissiveness may extend from the surface to the center of personality.

We should recognize two other dynamic considerations here that affect the general topic of social change and personality. Boundaries among all of the functions named above are both *permeable and flexible*. Some thoughts and affects, for example, may be conscious at one time and unconscious at others. Some aspects of our personalities we share with some other persons readily, and then slide them from the central to the peripheral region with other persons. At one time the author spoke with a teen-age girls' group on their psychosexual development. When the girls realized that this psychologist knew more about them in some ways than they knew about themselves, they began to listen and speak with freedom. The speaker became one of the girls among girls. Then one of the fathers came to get his daughter, entering the rear of the room very quietly and as unobtrusively as he could. But the spell was broken. The speaker was no longer one of the girls. The discussion continued, but the girls now concealed things that they had revealed when only "us girls" were present. What had been peripheral became central again. So, too, with other functions, the boundaries may be permeable.

From a sociological standpoint, the other consideration is important here: that is, that we may conceive of *external forces* impinging on the personality as so many vectors of force, moving the personality one way or another. If the forces approaching from different directions balance one another, the personality, as it were, goes nowhere. As is usually the situation, however, the forces do not balance one another, and the personality moves in a particular direction with a velocity that is the vector resultant of all the forces acting upon the personality.

Should anyone suppose that the personality is a helpless object moved willy-nilly by "external forces," he should pause to reconsider. The *"external* forces" are far and away only "external *forces"* insofar as *this particular personality* is concerned. Let the following incident illustrate: The author's oldest boy was with the Army in Viet Nam, then was removed to a Japanese hospital with some type of ear trouble in one ear. But this ear is his one good ear. He assured his parents and his wife that he would be all right. However, when his letters arrived, the news came as a tremendous blow to his wife and parents. The writer's son has a son who was two years old, and the writer has a son who was also two years old at the time. It is almost certain the soldier's wife discussed her fears and worries in front of her 2-year-old, and so did the parents in front of theirs. Did the same news make the same impact on these youngsters as on the adults? To ask the question is to answer it. What for one person is a long and powerful social vector, to another is scarcely a glancing blow. On the other hand, all personalities may be moved by some external forces of which they may or may not be aware, forces that they cannot interpret as powerful or weak for they are not even unconsciously cognizant of them, or being cognizant of them interpret them correctly as objectively overwhelming. The point is that most "external forces" are not of this latter variety, but are only *forces* impinging upon the personality because that particular personality makes them so.

Here we have the complicated picture of personality with which we have to deal in thinking of social change and personality: Functions of the personality may be divided up into the three Freudian constructs, id, ego, and superego, to which are added role functions, then into conscious and unconscious, into central and peripheral, into self (or selves) and non-self, and into multitudinous traits. These overlapping groups of functions are not static, for a particular function may slip from one category to another. And acting upon the whole personality are external forces, or vectors of force.

The parallel between the position described above and European stratification theory [see David and von Bracken (22)] is, the writer is inclined to believe, not fortuitous. Although developed independently, the concepts outlined here are not any more truly independent than is stratification theory; both have been influenced by Freudian theory, and more importantly both have been influenced by the raw data of personality. Since ancient times men have been forced to think of personality as having levels (what goes on in the "heart" is contrasted with more superficial, more readily observable behaviors). All the writer has done is to delineate more sharply among the levels that have been conceived at one time or another.

C. Social Changes in the Personality

What effect do social changes have upon the personality as conceived here? Just how far can social change change personality? A first answer may be given that reiterates a position stated before: Roles are interobject and interpersonal portions of the personality. *Therefore, any changes of roles are by definition social changes.* This principle cannot be emphasized strongly enough, for the man on the street and many scholars regard personality as a fixed entity inside the organism that somehow may be related to intelligence and attitudes and roles. This is not the view expressed here: personality *includes* intelligence, and *includes* attitudes, and *includes* roles. It is not just related to these functions; it *is* these functions among other things. Consequently, when the roles of an individual change, his personality changes. If he plays the role of a factory worker and he moves into the role of a supervisor, do not ask, "Has his personality *also* changed because his roles have changed?" for one cannot answer on these terms. Of course his personality has changed when his roles change, because *roles are part of the personality.* Insofar as roles may be changed, personality that consists in part of roles changes too. The degree of such change has to be determined by objective investigation. There has been some research along this line, but there has been mostly theorizing thus far [*cf.* Worchel and Byrne (89)].

Some of the reasoning employed in the preceding paragraph may seem circular: Personality is defined in a certain way, so that inevitably one must conclude that personality *does* change, because some social changes occur within the personality defined so as to include social functions. The only defense that can be given for such seeming circularity is that the concept of personality outlined above is a logical extension of interpersonal concepts of the personality that have arisen in the last few decades; and all the writer is doing is pointing out the implications of these concepts.

D. EXTERNAL FORCES AND PERSONALITY CHANGE

What about social changes that stem from what are identified above as external forces impinging upon the personality? Just how far can these alter the personality? Here let another general principle be stated: *roughly speaking, the closer to the center of the personality a function is, the more difficult is it for external forces to change it; and conversely, the farther from the center, the easier it is for external forces to change it.* By "closer to the center" and "farther from the center" the idea to be conveyed is that of progressing from id to roles, from unconsciousness to consciousness, from central to peripheral, and from self to not-self. Let us think, then, of how far social changes of an external nature can reach into and alter the personality.

Before doing so, a second possible *petitio principii* must be dealt with: If the functions of the personality were categorized into "closer to center" and "farther from center" on the basis of their vulnerability or invulnerability to social change, it would indeed be the case that the conclusion has been assumed in the premise. But such layers are not, at least substantially, defined by their responsiveness to social change. They are defined in some cases by their indispensability to the life processes, in others by their degree of concealment (conscious or unconscious), and in still others by the degree of investment or cathexis the individual has in them. While one cannot deny that in some instances the amount of resistance to social change is used as one criterion of centrality, in no case is the latter an essential criterion. Once again the circularity is only apparent. External social changes may be correlated with personality changes, but the latter are not necessarily defined by social changes.

1. *Prenatal Social Influences*

When one thinks of such social changes, one may be envisaging far-reaching sociological changes, poverty, unemployment, segregation and desegregation, mass migrations, war, and mass societal movements, such as the coming of Communism to Russia or Nazism to Germany. To be sure, these must be considered very seriously. But we begin much more modestly with the first social changes that occur in the life of an individual, those that take place in his intrauterine life. One may say the embryo or fetus has no personality—personality only comes after the child is born. Does it? Not according to the writer's definition: personality consists of the organized functions of the individual. Is an embryo an individual? Does he have organized functions? Then he is a personality.[2] By the end of the second month of gestation, when we usually say the embryo becomes a fetus, it is clear that there are individual differences among prenatal infants. But even before this time there are socially induced changes that affect the personality.

It cannot be considered trivial to say that social factors that could be different but that actually do occur make for personality changes even before the child is conceived. It is known now that malnutrition in the mother affects adversely the developing prenatal infant [see Mussen, Conger, and Kagan (59)]. Therefore, any social changes that generate conditions in which the mother's malnutrition reaches the point that her reserves are depleted also generate changes in the functioning of the child. We can think of a number of such social changes: wartime conditions, migration to barren places, general or specific famine conditions (especially if brought about by soil-depleting farming techniques), economic exploitation conditions, and so on. We have no way of knowing now just how much harm was done the unborn children of the Pilgrims in those hard months of their first implantation on the bleak New England scene. But we do know that socioreligious reasons moved the Pilgrims from the relative abundance of western Europe to this land. And if the results of controlled experiments on maternal malnutrition can be extrapolated to the Pilgrim mothers, we can only suppose some harm did occur to the infants conceived in those days.

---

[2] Since someone may say that the unborn child has life but not personality, a distinction must be made between life and personality [*cf.* the author's discussion on this very point (25)]. The concept of personality *in utero* does not depend on a particular definition of personality. Almost any definition implies the concept. That definition which across the years has proved the most acceptable of all (1) certainly does: "Personality is the dynamic organization within the individual of those psychophysical systems which determine his unique adjustment to his environment" (1, p. 48). At the very least, by the time the fetus has developed a rudimentary nervous system, "psychophysical systems" exist.

Contraction of Rubella or German measles in the first trimester of pregnancy can result in physical or mental defects in the child. Rubella is an epidemic disease: in other words, it is socially promulgated. The Rh factor, as is well-known, is an important genetically determined blood chemical or chemicals found in about 85 per cent of human beings. If an Rh-positive child is born to an Rh-negative mother, especially if the child is the second one, disastrous consequences to the child may result from the antibodies produced by the mother. But as with some *genetically determined* diseases, the mating of the mother and father comes about as a result of social factors, some purely accidental, some from societally determined assortative mating. If these two people or others so genetically constituted had not gotten together, the Rh incompatibility would not have occurred. So, too, if the Age of Tranquillizers had not been socially induced, several thousands of youngsters in western Europe would not now be known as thalidomide-deformed children. One could name a number of other purely physical agencies, with direct effect on the prenatal infant, that are easily seen to be as much the agents of a sociological change as they are of physics or chemistry.

Does the thesis, that the closer to the center the more difficult it is for external forces to change a personality function, hold in relation to prenatal socially induced personality changes? Generally, it appears to. The prenatal infant is initially primarily id, then very basic ego in his middle and later stages. The genes or the DNA molecules, genetic functions at the very center of the embryonic personality, seem to be the least amenable to change, but in this embryonic state of the individual where they are still close to the surface of the personality they are subject to mutations from various outside agencies, radiation in particular [see Dobzhansky (24); McClearn (55); and Sinnott, Dunn, and Dobzhansky (80)]. As controls develop in relation to genetically determined energies, controls that may be called the embryonic ego, *they* are more subject to these outside influences and the id processes less so.

Even if one could not support the thesis of increased difficulty in changing central embryonic and fetal processes, if we can grant that the prenatal infant either has or is a personality, the major point to press home is that that personality is liable to be changed and the subsequent effects noted when social forces are altered in such a way as to bring (usually indirect) influence to bear on the unborn child.

2. *Postnatal and Childhood Social Influences*

When we turn to the postnatal period, probably no one would deny that the young, even the newborn infant, has or is a personality. Here, on the one hand, someone may say the infinite plasticity of childhood is manifest. The tiny malleable creature can be molded, as John B. Watson asseverated many years ago, into anything the environment desires. Or on the other hand, another person may say that the infant is so close to being a gene-controlled creature, not much can be done to change him.

To the proponents of the first position it may be said that an ultra-environmentalist optimism must be tempered by the results of controlled research. As far as deep-lying id and ego processes are concerned, the social experiences of infancy seem not to be very potent. In the reviews of the research literature by Bettye Caldwell on the effects of infant care (19) and of Yarrow and Yarrow (91) on early experiences, it appears that the primary personality characteristics, the temperament patterns, are fairly well built in.

For example, the fact that a child has been breast- or bottle-fed, or given considerable or little oral gratification, or has been toilet trained one way or another, of itself seems to be of little influence in altering so-called oral and anal temperament characteristics. Shades of Freud and Abraham! Further, the more general patterns of differences among infants, such as irritability and activity level, can be predicted with a fair degree of accuracy from infancy to at least the early preschool years, even though infants and toddlers may have been subjected to a host of differing experiences (85).

In addition, one cannot ignore the comprehensive studies of the relative influences of genetic and environmental factors on intelligence functions that are fairly central to the personality.[3] The best studies utilizing monozygotic twins compared to dizygotic twins keep offering the figure of between 70 to

---

[3] The writer mentions this aspect because he knows the general interest in this area and not because he thinks it is of the *utmost* importance, important as it is. Kent Miller and the writer (29, 30) have tried to overcome the exceeding overemphasis on intelligence comparisons between Negroes and whites in the psychological literature by stressing temperament and social variables; but the writer does notice that the majority of citations to the 1960 *Psychological Bulletin* article refer to the section on intelligence.

80 per cent of the variance of intelligence test scores that may be attributed to heredity, and the remainder to environment [see Burt and Howard (17), and Nichols (45, 63)]. At least some of the specific factors of intellect, as well as general intelligence (Spearman's "g" factor), are subject to this same principle. In other words, social change as represented by altered social conditions, nutritional factors, differences in family and cultural patterns, and so forth have about 20 to 30 per cent of intelligence variance within which to work. These findings are in keeping with the thesis that the closer to the center of the personality a function is, the more difficult it is to change it.

In reference to the proponents of the second position—that the infant is so close to being gene-controlled that he can scarcely be changed at all—it may be stated that later developed functions, and especially those that do not become deeply buried by repressions, may yield to social changes to a marked degree, so that even relatively early vast personality alterations may take place [even in intellectual functions, if results of recent rat experiments with environmental impoverishment or complexity may be extrapolated to man (70), though these results challenge some previous studies]. These personality alterations may be, and in all likelihood usually are, the resultant of combined internal and external changes. But there is no requirement that social change has to be the sole determinant of a personality change, but may indeed be the precipitator or catalyst of such change, a necessary even if not a sufficient condition.

In this connection, as is often the case, reference to extremes of behavior, as in abnormal states of personality, clarifies a process that may not be seen as clearly in the less pronounced normal state. Anna Freud, Sigmund Freud's exceptionally brilliant daughter, has recently called attention to a phenomenon that proved distressing to psychoanalysts when they began longitudinal investigations of children. Some similarity between adult and childhood neuroses can be found in both symptoms and dynamics. But among the disappointments that analysts met

> was the realization that in spite of all the links between infantile and adult neuroses, there is no certainty that a particular type of infantile neurosis will prove to be the forerunner of the same type of adult neurosis. On the contrary, there is much clinical evidence which points in the opposite direction. For example in the unmanageable state of a four-year-old which equals in many ways that of a juvenile or adult delinquent insofar as both give free rein to their impulses, especially the aggressive ones, and both attack, destroy, and appropriate what they desire without regard for other people's feelings. For all this similarity, the early delinquent behavior need not turn later into a true delinquent

> cases; the child in question may develop into an obsessional character or obsessional neurotic rather than into a delinquent or criminal. Many children who begin with a phobia or anxiety hysteria grow later into true obsessionals. Many with early obsessional symptoms such as washing compulsions, ritualistic touching, arranging of details, etc., who resemble adult obsessionals in every way while they are young, are nevertheless predestined to develop in later life not obsessional neurosis but schizoid and schizophrenic states instead (35, p. 150-151).

Anna Freud attributes these fairly fundamental changes to id and deep-lying ego operations. One would not wish to minimize these internal changes that result in such dramatic overtly observable, but deep, personality changes. Yet it is well to suggest also that the external world does not remain static while the child is maturing, so that a number of the personality changes to which Anna Freud refers may well stem in part from the child's changing world.

Yarrow and Yarrow (93) point out, for instance, that family structure and dynamics do not remain the same across the years. Every clinician with fairly extensive experience has probably had a similar experience to one of the author's. The clinic he headed dealt some years ago with a very bright boy of 9 with whom the author had the major therapeutic contacts. In fact, this boy was one for whom the clinic had to initiate activity group therapy, for play therapy was too infantile for him. Some years later the writer met the mother who told of the boy's quite substantial recent accomplishments. The writer said with some pride, "I think I can take credit for some of that." The mother paused for a moment, then candidly remarked, "I don't think that had a great deal to do with it. What I learned from the clinic about handling him I'm sure had a great deal more to do with it." Deflated, one therapist's ego balloon! While it is not necessary to admit that the carefully devised activity therapy had no effect, one can readily recognize that the change in child management practices in that home probably did have more to do with altering the boy's personality than did the therapy.

Not long ago a case was terminated with a 13½-year-old girl who had to be carried bodily for her second session from her car to the playroom, not because she could not walk but because she balked and locked herself in the car. Screaming and struggling all the way, she yielded only to superior force. Inside the playroom she made the welkin roar with her screams and high-pitched cursing. She used language on the therapist that her mother did not know she knew. The least rabid of her epithets was "Stinker!" and "Stupid!" For a number of sessions she continued her tantrums and tirades. Some sessions later, however, this same sweet young lady had arrived at the point where she could express liking for the therapist, adding in one period, as she looked up a bit shyly, "I even love you." She had repressed the knowledge that she loved the therapist by the next session; but after a somewhat stormy hour resulting from several factors, she made one pipe-cleaner figure for herself and one for the therapist. The changes that occurred in this girl—not limited by any means to the therapy hour—were not solely dependent upon her therapy contacts. Her parents drastically altered their management techniques as a result of clinical guidance, including the use of Ginott's *Between Parent and Child* (40). Family structure and dynamics may change, altering the personalities of children to a measurable degree.

Less dramatically, usually, but nonetheless effectively, for better or for worse nonclinical families change and personalities of youngsters change along with them. How much personality change results from so much change in the family *qua* family can only be conjectured at this time. Maturation effects are sometimes confounded with learning effects, as we may infer from experiments on "imprinting" [see Hess (42)], so that we may think something results from social factors when it is really the result of maturing functions. Yet even these could not mature without an appropriate environment.

The age and stage at which certain social experiences occur may make considerable difference as to their effectiveness in altering personality. Separation from parents has been taken as one very traumatic social experience that is supposed to alter a young child's personality profoundly. Yet Yarrow (90) has shown from going over the studies in this field that the effects of separation *per se* may not be traumatic for the very young infant, say in the first six months of postuterine existence, and even afterward may be more determined by events subsequent to the separation, such as being placed in an institution. However, some evidence exists for the conclusion that immediately observable behavior in the infant following a separation experience may not be a good indicator of the degree of disturbance that shows up later. Later separation at certain stages may be more disastrous than at others. A case known to the writer is that of a girl of 5½ whose mother died in the midst of the girl's partly expressed oedipal fantasies. The tremendous guilt aroused by the fulfillment of these fantasies of killing mother and marrying daddy was overpowering, so that profoundly disturbing patterns of affect and behavior showed up.

Whatever the directly traceable effects of early separation are, we have to keep in mind from the standpoint of our inquiry here *any* changes, including changes in roles of the child and his important others, prior to, accompanying, or subsequent to the parental separation itself are *social* changes; and whether their effects are equal to, or less than, or more than the actual separation, it is the social changes *of some kind* that are or bring about beneficial or deleterious personality changes, possibly reaching fairly deeply into the central portions of the personality.

3. *Social Influences in Early Adolescence*

Space does not permit further elaboration of the impact of social forces on childhood personality, or of their effects on the aging personality. But it is necessary to take time to discuss changes surrounding puberty and adolescence. Psychoanalysts (14, 37, 49), social psychologists (79), child psychologists (54), including Piaget and his followers (35), regard the beginning of the adolescent period as somewhat of a crisis, at least in our society. McCandless calls it "the fifth psychosocial crisis." Sherif and Cantril are close to psychoanalysis when they speak of the "re-formation of the ego in adolescence." Although there may be an overemphasis on adolescence as the period of *Sturm und Drang,* storm and stress, nevertheless it has been recognized for a long time that adolescence has been a crisis period in the West.

Some evidence has accumulated that in our era society has changed enough to make the period even more stressful yet. Perhaps most potent is "the American sex revolution," as Sorokin (83) dubbed it. Probably, as a survey by the *National Observer* not long ago indicated, not much more premarital sexual freedom actually is present now than there was a generation ago [see Klemer (46)]. However, *attitudes* toward sex have changed in many ways [see Schur (75)]. And adolescents have to face a world that more than ever gives them equivocal guidance in sexual matters. Other forces of major magnitude impinge on the lives of young adolescents, such as a plethora of violence, mayhem, and murder on TV and movie screens, large scale financial crimes in high places, revolts against exploitation of oppressed peoples, and issues of peace and war brought to focus by Viet Nam.

Do these vast social changes alter the personalities of youth in the crises of puberty and early adolescence? One could cite opinion studies and questionnaire responses of youth relating to specific opinions and attitudes on one or another issue, studies that do indicate changes. And insofar as opinions and attitudes are personality functions, it is evident that youthful personalities have changed. However, deep changes in the adolescent personality probably do not occur primarily because of the vast social changes going on about the pubertal and early adolescent individual.

If we turn to recent intensive studies of adolescent personality undertaken by psychoanalysts (11, 14, 41, 49), which have been carried out in the present era, in the midst of all the social upheavals we find that they tend to support the long-held views of psychoanalysis (34) that the external world, while surely playing a part in the pubertal crisis, is not as important as the intrapsychic (the more central) functions in shaping the crises of this stage. Further, Quay and Quay (67), working with seventh and eighth graders, repeated the work of Peterson (64) that delineated the problem behaviors of middle childhood; they produced by factor analysis the same two factors in these pubertal youngsters that Peterson had found in latency age children and labelled "Conduct Problems" and "Personality Problems." From a practical, clinical standpoint this division into only two factors is nearly useless. But when Quay and Quay found the same two factors in over 500 seventh- and eighth-grade children, and only one other weak factor apparently related to behavioral immaturity in the eighth graders, from a theoretical standpoint the results suggest that even today the pubertal child has not changed a great deal in his problem orientation from what he was in the prepubertal period. In other words, the storms of the present day have not inundated the pubertal youngster in his personality crisis period. Similar findings were obtained by the Behavioral Classification Project (26, 28, 32). The writer and associates discovered that among the many factors of child behavior, principally behavior problems, which they found among hundreds of clinical and normal children, only one, which was a specifically maturity-immaturity factor, was age-related among children 6 through 13 years of age.

The theory, then, that the tremendous social changes of today's rapidly moving world are not reaching the core of the young adolescent's personality with the devastating impact they might seem to have, is not groundless [see Klemer (46)]. That one of America's most respected authorities on family life agrees with this theory is support of an indirect nature. Blaine R. Porter writes of the present teen-ager, "Some four or five per cent (or even 10 per cent) get into trouble, and the trouble they get into is extremely serious. But considering the temptations, frustrations, and complexities to which they are subjected, and the confusion that surrounds them, the 90-95 per cent who do not get into trouble deserve congratulations" (65, p. 147).

4. *Social Influences in Later Adolescence*

In the latter years of adolescence more and more youth in our society attend college. From what one reads by alarmed adults in papers and magazines, atheistic and communistic professors are influencing young people toward evil ways. There must be some magical attribution of power to college professors at this point, for the research—to say nothing of a professor's chastening experience—does not bear out the vast influence ascribed to professors, either for good or ill. It is not that under college suasions changes do not occur, sometimes changes of a fairly dramatic nature, but it would seem that they do not ordinarily reach to the depths of the young person's personality.

Without any scientific study it can be seen that new roles are assumed when a young person takes up a college career; and older, established roles are altered, like those of being a son or a daughter, a male or a female, and being an adolescent. As argued previously, changes in roles *are* changes in personality. Attitudes of some sorts are changed, partly by the impact of the teachers of a college or university, but mostly by the general atmosphere of the campus. A study colleagues and the writer carried out about a dozen years ago (31) showed that dating attitudes and practices may change radically in the first three months of college life. Other changes of this nature do take place, so that the critics of higher education may confuse these changes with the more fundamental ones they think are taking place. The more fundamental functions, however, probably are not altered.

There has been, to be sure, a consistent showing that students in college become increasingly liberal in their religious thinking, decreasingly ethnocentric, and increasingly independent and unconventional [see Lehmann *et al.* (50), Prien (66), Sanford (71), and Webster (86)].

But even if these findings of change do mean that students are influenced toward more liberal *expressed* attitudes, it seems reasonable to believe that the inner core of the personality in which the beliefs are rooted may not change as a result of external social changes as much as might appear. One study of a generation ago (39), for example, runs counter to the general trend, finding that in several colleges students did *not* change in the direction of their professor's religious beliefs. And in the present decade Nichols (62) has carried out a four-year longitudinal study with three temperament scales on the National Merit Scholars in a large number of different colleges. As in previous studies, colleges are shown to be somewhat different from one another; and different curricula seem to have some differential effects. Some few temperament changes appear. But in general there has not been a great deal of change effected in the basic personality structure of students. Nichols concludes that the major portion of the differences found in his subjects could be attributed to events that happened before they entered college.

The student who has been a "hard shell conservative" seems to become what one minister friend calls a "hard shell liberal." Some students who adopt a liberal outlook in college may be only conservatives in disguise, for research (48) shows that the farther out of college many a man gets, the more conservative his views become.[4] The writer is reminded of a member of a faculty committee who was studying the drinking problems on his campus. This man said in one of the meetings, "My views are changing rapidly; I have a daughter ready to start in the university next year." Observation of the man's general behavior in the committee would lead one to believe that his liberal views on drinking have been overlaid on some much deeper convictions about the wrongs of drinking, views now beginning to reassert themselves. One cannot forget in assessing the impact of college on personality that the majority of Southern legislators, both national and state, who have been so consistently antiliberal in their legislative actions, are college-trained men who for four years at least were exposed to many influences that should have moved them toward liberal social attitudes.

E. Central Personality Changes

Having said all these things about what cannot be changed by ordinary social changes in the development of the personality, to keep perspective we must hasten to reaffirm that such changes *can* occur under certain circumstances on a very deep level of the personality. Such changes are not impossible, they are only more difficult. Briefly, some of the circumstances and some of the changes that occur may be illustrated by the following.

1. *Facilitated Behaviors*

There are first what may be called "facilitated behaviors," as opposed to stress-induced behaviors. So much has been written about stress and its effects that we tend to overlook positive forces that may penetrate to the deeper levels of the personality and profoundly alter strata far from the surface. A book may reach the depths of one person, a sunset another, a mountain or canyon another, an opera or a person another. In all these the individual may be able to say "from the depths of his heart" that he has never been the same since "that experience." Granted that "that experience" (either one time or prolonged) is probably only catalytic, acting upon a concatenation of forces already in existence, the reality of the realignment of certain temperament patterns can only be doubted by challenging most of the evidence in many cases. Out of the many possible "facilitated behaviors" that could be discussed here, only a very limited sample can be taken.

---

[4] Studies of persons 14 or 15 years out of college contradict the general findings of increased conservatism or interpret conservative trends or liberal trends as primarily a function of the changing times [see Bender (10), and Nelson (60)]. It is true that waves of conservatism and liberalism move across the land; one expects, then, that individuals will tend to go along with the times, whether they are in college or out. That one study (Nelson) shows a liberal trend from 1936 to 1950 and the other (Bender) a conservative trend from 1940 to 1956 should not be surprising (the McCarthy era was in the early 1950's). The main point of the text discussion is that college does not make the radical changes in personality that its friends and enemies hope or fear it does. There is no need to deny, or rather there is need to assert, that social changes do affect personality; but the superposition of college influence on the main societal trends seems to make not a great deal of difference. Relatively small shifts are reported in the study cited showing the liberal trend, and somewhat larger shifts in the other study showing conservative trends.

For illustrative purposes, what happens under the influence of psychotherapy is important, not merely because some psychologists, psychiatrists, social workers, and others have invested a great deal of their lives in psychotherapy, and hosts of clients or patients have committed their destinies to it, but also because psychotherapy represents in miniature what facilitation if any can occur as a result of deliberately manipulating certain interpersonal variables in order to bring about more or less deep-lying changes "for the better" in the personality.

It is difficult to assess the changes that may be brought about by psychotherapy (18, 52) in the underlying patterns of the personality. For many years it has seemed that psychotherapy cannot be much good in altering the personality, for the conclusions reached by Eysenck (33) in the early part of the fifties have stood unchallenged until recently. Eysenck surveyed the studies on psychotherapy up to that time and concluded that about two-thirds of those who had psychotherapy recovered, and that of similar persons who had not had psychotherapy about two-thirds recovered. More recently, however, a serious challenge (45) has been made to Eysenck's position that gives hope to those who engage in psychotherapy. They have felt all along that *something* of a more or less profound nature was happening to their clients in response to the social changes they were providing in the form of psychotherapy.

Of course, controlled research is the only way to ascertain whether the psychotherapist's feelings are justified or not. Every palmist and astrologer can point to some rather astounding "successes" achieved with their clients. Further, what sometimes seems like profound change may only be reduction in anxiety to a marked degree. On going over his records of cases he had seen in the last two years before leaving one university, the author discovered to his dismay that three young men, whose cases had been terminated on the basis of the usual signs employed in psychotherapy, had shown by their subsequent behavior on campus not to have been altered basically at all by therapy. Where before they had displayed psychopathic trends but had come for therapy because of anxieties they were experiencing, now they were displaying the same psychopathic trends without the anxiety.

Nevertheless, it does seem that some of the people who have undergone psychotherapy have had fairly central alterations in their personalities, if not of basic id processes, at least of deeper ego and other unconscious functions (36). Consider, for example, the woman who "had suffered many things of many physicians," having gone to a dozen physicians and renowned clinics with her psychosomatic problems; in three months' psychotherapy she found more help of a lasting nature than from all the physical therapies. Other cases from the files of other psychotherapists, complete with follow-ups, would demonstrate the same thing. As is true of any social process intended to facilitate behavior, it must be said of psychotherapy: that profound change does not occur in every instance does not mean it has not occurred in some instances.

Very closely related to psychotherapy, and indeed utilized by many psychotherapists, is hypnosis, which may be subsumed here under facilitating factors. Although the authorities who have written on hypnosis are legion, by far the most persistent researcher and critical authority in the field is Barber and his associates, with about half a hundred experimental and theoretical articles on the subject. The most salient publications (6, 7, 8, 9) dispel many myths about hypnosis. With some of Barber's contentions other authorities may differ,[5] but a number of his and others' experiments lead very strongly to several surprising conclusions: First, "hypnosis" itself is an unclear concept, sometimes meaning an independent variable, sometimes a dependent variable, and offtimes some unspecified combinations. Second, many subjects perform "hypnotic" behaviors without benefit of induction procedures. Third, the prior relation to the hypnotist seems to be very important in many cases. And fourth, the prior personality structure of the subject, and in particular his "suggestibility," appears to be a determining factor—whether they are "hypnotized" or not, some persons perform the feats of a presumably hypnotized individual at the behest of an experimenter. There are other conclusions to be drawn from Barber's work, but these are most pertinent to our quest.

If the foregoing conclusions are correct, "hypnosis" would seem to be a variety of interpersonal relation that would serve as a social change much the same as does psychotherapy. As far as the writer knows, despite many studies of "posthypnotic suggestion," dealing primarily with retention of specified memory content, no studies have been made of long-term effects on personality organization, and especially any that might bear on temperament reorganization. Presumably, however, given a subject who is suggestible enough, a therapist with prestige yet with whom the individual relates well, and a continued "hypnotic" situation intermittently repeated long enough, we should expect far-reaching changes in the personality (to some extent in each of the dyad). Hypnoanalysis [see Arluck (4), and Wolberg (87, 88)] contains most if not all these ingredients; its proponents claim just such fundamental alterations in the personality as suggested above. As for drastic changes from one or a few sessions, the writer supposes ideal conditions of preparation in the subject's life would be needed—not an impossible condition, but unlikely in most cases.

---

[5] Sarbin (72) some time ago had interpreted hypnotic behavior from a role theory standpoint. His conclusions appear to be similar to Barber's, especially as to the subject's relation to the hypnotist. Rotter, in discussing the role of the psychological situation from his Social Learning Theory standpoint (68), develops concepts that reinforce the interpretation of hypnosis as described by Barber.

As stated previously, other social experiences too numerous to mention may bring about substantial deep-level change in the personality. Whatever happens, for example, in the understanding acceptance of a teacher's after-school conference with a pupil, may "make all the difference in the world" to the student. And what happens in a ministers' study may likewise produce such changes. The writer spent a number of years in the ministry. One of his most vivid memories of his pastoral work is of one morning during the Second World War when a professional man, not a member of his church but more often in attendance than a good many members, called. He said in an urgent tone, "Ralph, can I see you right away?" When he came into the study, he dropped into the chair across from the minister and said, "The only thing I can think of is that I want to kill myself." He went on, "I have everything to live for. I have a lovely wife and daughters. I have everything to live for." He repeated that same thought several times during the conversation. Since that time the writer has come to realize the significance of the repetition of this idea, but at the time of this story he did not understand it. The distressed man spoke of the band like an iron band he felt about his chest, and how he could not breathe because the band was tightening on him so. After the two had talked for a short while, the minister said, "Paul"—that is not his name—"Paul, let us pray about this." So they knelt down at their chairs and the writer prayed briefly. Then the other began to pray. At first his words came out haltingly. Then he began to cry and he said, "I feel the band is loosening about my chest. I feel I can breathe again." In a short while he rose, they shook hands, and he left. It is not entirely clear just what happened that day. All that we do know is that this man was ready to take his own life, and that now nearly a quarter of a century later he has not done so. Possibly no *basic* alteration in this man's personality structure took place in response to introducing certain social change, but something fairly profound, at a deep, unconscious level, did take place.

2. *Stress-Induced Behaviors*

Just as "facilitated behavior" may be brought about by beneficial changes in society, so, too, central changes may be brought about by stress. Most biological and social scientists are familiar with Selye's (76, 77, 78) "general-adaptation syndrome." If a stress is prolonged sufficiently, the organism reaches the exhaustion stage and expires. This would seem to be a fairly extensive alteration in the personality. Such stresses may be physical, such as extreme cold or heat, or infection or pain. Even these may be the result of social processes, as in a war situation or a concentration camp. In fact, much that we know about reactions to stress came from studies made during wartime (82) and reports of former inmates of concentration camps (69). The results of deliberate attempts to alter the basic personality structures of prisoners of war, as in China's so-called "brain-washing" of American prisoners in Korea, are somewhat in doubt (44); but enough evidence has accumulated to recognize that some prisoners were deeply affected. Indeed, just being a prisoner of war may wreak profound havoc (20, 21). Likewise, although some returnees from concentration camps seemed not to have been deeply changed (3), the preponderance of evidence (12, 15, 53) suggests that the extreme stresses set up by a society bent on mass destruction,[6] stresses both physiological, psychological, and sociological, wrought alterations in central portions of personalities that can probably never be effaced in those who survived.

In recent years some investigators have been coming to realize the deep-lying effects a system of segregation can produce in the personalities of those who are oppressed by it as well as in those who are the oppressors, or what changes may come when desegregation takes place. As one of the bishops of the Methodist Church in India said after he, an outcast, had become a Christian, "One of the most important things my becoming a Christian has meant to me is that I can meet a man and look him in the eye." We have to remember that slavery and segregation in this country were introduced as social changes and that the very depths of personalities were reached by these changes, so that, for example, both many whites and many Negroes came to feel that anything different would be immoral. Young Negroes who have been growing up in a different day may not understand [though some may, see Solomon and Fishman (81)] that some of the opposition—in fact, the most basic opposition—their elders manifested at the start of the Negro Revolution was not from fear of consequences, but from a conviction reaching deep into their unconscious minds that what their youth were doing is immoral. The same holds for a large number of white persons. Their antipathy or apparent indifference to the Negro's need for freedom may not stem from sadism or delight in having the fruits of exploitation, but from the feeling that associating with Negroes on an equal plane is just downright immoral. How far the social changes of the present will reach these underlying superego attitudes is a moot question, although some indications are that changes can be wrought (13).

---

[6] This is a society organized for a heinous purpose. A disorganized society may likewise achieve evil ends [see Blumer (16)].

F. POSTSCRIPT

One last word about social change and personality: we are not at all sure how social change does bring about personality change, although some investigators have been theorizing for years about the matter, as Freud, Adler, and Jung all did.[7] A more modern attempt (38) leaves the writer rather cold. When observable social changes are followed by observable personality changes (57), we may tend to think we have at least identified the causal relations. But as scientists have been warned again and again, correlation does not mean causation. What about the times when vast social change is not followed by the expected personality change (43)? And what about the vast personality change not preceded by any observable social change?

An instance of the latter may serve to bring humility to the ranks of those who are presumed to be authorities in the area of personality and social psychology. The man was a rancher friend with a fair-sized family. But he drank extremely heavily. In fact, he told the writer—and there was independent evidence as to the correctness of his figure—that he spent two hundred dollars a month on liquor for himself, not to entertain, but just for himself. And that was under price control. The figure would be at least double today. When this man's wife had a very unfortunate accident, the writer thought it might bring the friend to himself. But, no. Then his son had a serious accident, a truly horrible thing. Surely, the writer thought, the man would be reached. But, no, he kept on with his drinking. The exhortations of friends and even the help of others during his very trying times seemed to be of no avail. A year after the writer's family had moved from the community, the man quit drinking—just like that. No tapering off, no sanitarium "drying out," no religious conversion, no AA, nothing that could be seen—he just quit. What happened? The writer can only hazard some guesses about the cumulative effects of a string of misfortunes. But actually nothing is really known. In some ways we are not too far from that situation in our scientific position in respect to social change and personality in general. We are convinced that social change brings personality change. The writer is convinced, as are many, that the more central the personality function is, the more difficult it is to change.

---

[7] The latest has just come to hand, N. Sanford's *Self and Society: Social Change and Individual Change* (New York: Atherton, 1966), whose aim is to provide a basis for planned action affecting individuals and groups.

REFERENCES

1. ALLPORT, G. W. Personality: A Psychological Interpretation. New York: Holt, 1937.
2. ———. Traits revisited. *Amer. Psychol.,* 1966, **21**, 1-10.
3. ALLPORT, G. W., BRUNNER, J. S., & JANDORF, E. M. Personality under social catastrophe: Ninety life histories of the Nazi revolution. *Charac. Personal.,* 1941, **10**, 1-22.
4. ARLUCK, E. W. Hypnoanalysis: A Case Study. New York: Random House, 1964.
5. ASCH, S. E. Studies in the principles of judgments and attitudes: II. Determination of judgments by group and by ego standards. *J. Soc. Psychol.,* 1940, **12**, 433-465.
6. BARBER, T. X. Antisocial and criminal acts induced by "hypnosis." *Arch. Gen. Psychol.,* 1961, **5**, 301-312.
7. ———. Hypnotizability, suggestibility, and personality: V. A critical review of research findings. *Psychol. Rep.,* 1964, **14** (Monograph Supplement 3-V14), 299-320.
8. ———. Experimental analyses of "hypnotic" behavior: A review of recent empirical findings. *J. Abn. & Soc. Psychol.,* 1965, **70**, 132-154.
9. BARBER, T. X., & CALVERLY, D. S. Empirical evidence for a theory of hypnotic behavior: Effects on suggestibility of five variables typically included in hypnotic induction procedures. *J. Consult. Psychol.,* 1965, **29**, 98-107.
10. BENDER, I. E. Changes in religious interest: A retest after 15 years. *J. Abn. & Soc. Psychol.,* 1958, **57**, 41-46.
11. BERGEN, M. A. Some observations of maturational factors in young children and adolescents. In R. S. Eissler (Ed.), *The Psychoanalytic Study of the Child (Vol. XIX)*. New York: Internat. Univ. Press, 1964.
12. BETTELHEIM, B. Individual and mass behavior in extreme situations. *J. Abn. & Soc. Psychol.,* 1943, **38**, 417-452.
13. BETTELHEIM, B., & JANOWITZ, M. Social Change and Prejudice. New York: Free Press of Glencoe, 1964.
14. BLOS, P. The initial stage of male adolescence. In R. S. Eissler (Ed.), *The Psychoanalytic Study of the Child (Vol. XX)*. New York: Internat. Univ. Press, 1965.
15. BLUHM, H. O. How did they survive? Mechanisms of defense in Nazi concentration camps. *Amer. J. Psychother.,* 1948, **2**, 3-32.
16. BLUMER, H. Social disorganization and individual disorganization. *Amer. J. Sociol.,* 1937, **42**, 871-877.
17. BURT, C., & HOWARD, M. The relative influence of heredity and environment. *Brit. J. Stat. Psychol.,* 1957, **10**, 98-104.
18. BYRNE, J. Assessing personality variables and their alterations. In P. Worchel & D. Byrne (Eds.), *Personality Change.* New York: Wiley, 1964.
19. CALDWELL, B. M. The effects of infant care. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of Child Development Research (Vol. 1)*. New York: Russell Sage Foundation, 1964.

20. Curle, A. Transitional communities and social reconnection: A follow-up study of the civil resettlement of British prisoners of war, Part I. *Hum. Rel.*, 1947, **1**, 42-68.

21. Curle, A., & Trist, E. L. Transitional communities and social resettlement of British prisoners of war, Part II. *Hum. Rel.*, 1947, **1**, 240-288.

22. David, H. P., & von Bracken, H. (Eds.). Perspectives in Personality Theory. New York: Basic Books, 1957.

23. Doby, J. T. Introduction to Social Psychology. New York: Appleton-Century-Crofts, 1966.

24. Dobzhansky, T. Mankind Evolving: The Evolution of the Human Species. New Haven, Conn.: Yale Univ. Press, 1962.

25. Dreger, R. M. Fundamentals of Personality: A Functional Psychology of Personality. Philadelphia: Lippincott, 1962.

26. ———. A progress report on a factor analytic approach to classification in child psychiatry. *Psychiat. Res. Rep.*, 1964, No. **18**, 22-58.

27. ———. Custom and conscience functions of the superego. Unpublished paper, Louisiana State University, Baton Rouge, Louisiana, 1966.

28. Dreger, R. M., & Dreger, G. E. Proceedings of the Technical Assistance Project held at Jacksonville University, August 16-17, 1962. Behavioral Classification Project, Report No. 1. Jacksonville, Fla.: Behavioral Classification Project, 1962.

29. Dreger, R. M., & Miller, K. S. Comparative psychological studies of Negroes and whites in the United States. *Psychol. Bull.*, 1960, **57**, 361-402.

30. ———. Recent research in psychological comparisons of Negroes and whites in the United States. Paper read at Southeastern Psychological Association, Atlanta, Georgia, April, 1965.

31. Dreger, R. M., Smith, W. D., & Wieland, R. G. Changes in dating attitudes and practices in the first three months of college life. Unpublished paper, Florida State University, Tallahassee, Florida, 1954.

32. Dreger, R. M., Reid, M. P., Lewis, P. M., Overlade, D. C., Rich, T. A., Taffel, C., Miller, K. S., & Flemming, E. L. Behavioral Classification Project. *J. Consult. Psychol.*, 1964, **28**, 1-13.

33. Eysenck, H. J. The effects of psychotherapy: An evaluation. *J. Consult. Psychol.*, 1952, **16**, 319-323.

34. Fenichel, O. The Psychoanalytic Theory of Neurosis. New York: Norton, 1945.

35. Flavell, J. H. The Developmental Psychology of Jean Piaget. Princeton, N. J.: Van Nostrand, 1963.

36. Frank, J. D. Persuasion and Healing. New York: Schocken, 1963.

37. Freud, A. Normality and Pathology in Childhood: Assessments of Development. New York: Internat. Univ. Press, 1965.

38. Gendlin, E. T. A theory of personality change. In P. Worchel and D. Byrne (Eds.), *Personality Change*. New York: Wiley, 1964.

39. Gilliland, A. R. The attitude of college students toward God and the church. *J. Soc. Psychol.*, 1940, **11**, 11-18.

40. Ginott, H. *Between Parent and Child*. New York: Macmillan, 1965.

41. Gyomroi, E. L. The analysis of a young concentration camp victim. In R. S. Eissler (Ed.), *The Psychoanalytic Study of the Child (Vol. XVIII)*. New York: Internat. Univ. Press, 1963.

42. Hess, E. H. Two conditions limiting critical age for imprinting. *J. Comp. & Physiol. Psychol.*, 1959, **52**, 515-518.

43. Hilberg, A. The Destruction of the European Jews. Chicago, Ill.: Quadrangle, 1961.

44. Holt, R. R. Forcible indoctrination and personality change. In P. Worchel & D. Byrne (Eds.), *Personality Change*. New York: Wiley, 1964.

45. Kiesler, D. J. Some myths of psychotherapy research and the search for a paradigm. *Psychol. Bull.*, 1966, **65**, 110-136.

46. Klemer, R. H. Student attitudes toward guidance in sexual morality. In E. M. Schur (Ed.), *The Family and the Sexual Revolution, Selected Readings.* Bloomington, Ind.: Indiana Univ. Press, 1964.

47. Kluckhohn, C., Murray, H. A., & Schneider, D. M. (Eds.). *Personality in Nature, Society, and Culture.* New York: Knopf, 1953.

48. Kuhlen, R. G. Personality change with age. In P. Worchel & D. Byrne (Eds.), *Personality Change.* New York: Wiley, 1964.

49. Laufer, M. Ego ideal and pseudo ego ideal in adolescence. In R. S. Eissler (Ed.), *The Psychoanalytic Study of the Child (Vol. XIX).* New York: Internat. Univ. Press, 1964.

50. Lehmann, I. J., Sinba, B. K., & Harnett, R. T. Changes in attitudes and values associated with college attendance. *J. Educ. Psychol.*, 1966, **57**, 89-98.

51. Lewin, K. Some socio-psychological differences between the United States and Germany. *Charac. & Personal.*, 1936, **4**, 265-293.

52. Luborsky, L., & Schimek, J. Psychoanalytic theories of therapeutic and developmental change: Implications for assessment. In P. Worchel & D. Byrne (Eds.), *Personality Change.* New York: Wiley, 1964.

53. Luchterhand, E. Survival in the concentration camp: An individual or a group phenomenon? In B. Rosenberg, I. Gerver, & F. W. Howton (Eds.), *Mass Society in Crisis: Social Problems and Social Pathology.* New York: Macmillan, 1964.

54. McCandless, B. R. Children and Adolescents: Behavior and Development. New York: Holt, Rinehart & Winston, 1961.

55. McClearn, G. E. Genetics and behavior development. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of Child Development Research* (*Vol. 1*). New York: Russell Sage Foundation, 1964.

56. Mead, G. H. Mind, self, and society from the standpoint of social Behaviorist (ed. by C. N. Morris). Chicago, Ill.: Univ. Chicago Press, 1934.

57. Merton, R. K. Social structure and anomie. *Amer. Sociolog. Rev.*, 1938, **3**, 672-682.

58. Murphy, G. Personality: A Biosocial Approach to Origins and Structure. New York: Harper, 1947.

59. Mussen, P. H., Conger, J. J., & Kagan, J. Child Development and Personality (2nd ed.). New York: Harper & Row, 1963.

60. Nelson, E. N. P. Persistence of attitudes of college students fourteen years later. *Psychol. Monog.*, 1954, **68** (2), Whole No. 373.

61. Nichols, R. C. The inheritance of general and specific ability. National Merit Scholarship Corporation, Research Reports, Vol. 1, No. 1, Evanston, Illinois, 1965.

62. ———. Personality change and the college. National Merit Scholarship Corporation Research Reports, Vol. 1, No. 2, Evanston, Illinois, 1965.

63. ———. The National Merit twin study. In S. G. Vandenberg (Ed.), *Methods and Goals in Human Behavior Genetics.* New York: Academic Press, 1965.

64. Peterson, D. R. Behavior problems of middle childhood. *J. Consult. Psychol.*, 1961, **25**, 205-209.

65. Porter, B. R. American teen-agers of the 1960's—Our despair or hope? *J. Marriage & Family*, 1965, **27**, 139-147.

66. Prien, E. P. Personality correlates and changes in proworldmindedness and antiworldmindedness following an intercultural experience. *J. Soc. Psychol.*, 1966, **68**, 243-247.

67. Quay, H. C., & Quay, L. C. Behavior problems in early adolescence. *Child Devel.*, 1965, **36**, 215-220.

68. ROTTER, J. B. The role of the psychological situation in determining the direction of human behavior. In R. C. Teevan & R. C. Birney (Eds.), *Theories of Motivation in Personality and Social Psychology.* New York: Van Nostrand, 1964.

69. ROSENBERG, B., GERVER, I., & HOWTON, F. W. (Eds.). *Mass Society in Crisis: Social Problems and Social Pathology.* New York: Macmillan, 1964.

70. ROSENZWEIG, M. R. Environmental complexity, cerebral change, and behavior. *Amer. Psychol.,* 1966, **21**, 321-332.

71. SANFORD, N. (Ed.). Personality development during the college years. *J. Soc. Issues,* 1956, **12**, No. 4.

72. SARBIN, T. R. Contribution to role-taking theory: I. Hypnotic behavior. *Psychol. Rev.,* 1950, **57**, 255-270.

73. ———. Role theory. In G. Lindzey (Ed.), *Handbook of Social Psychology, Vol. I. Theory and Method.* Cambridge, Mass.: Addison-Wesley, 1954.

74. ———. Role theoretical interpretation of psychological change. In P. Worchel & D. Byrne (Eds.), *Personality Change.* New York: Wiley, 1964.

75. SCHUR, E. M. (Ed.). The Family and the Sexual Revolution, Selected Readings. Bloomington, Ind.: Indiana Univ. Press, 1964.

76. SELYE, H. A syndrome produced by diverse nocuous agents. *Nature,* 1936, **138**, 32.

77. ———. The Physiology and Pathology of Exposure to Stress; a Treatise Based on the Concepts of the General-Adaptation-Syndrome and the Diseases of Adaptation. Montreal: Acta, 1950.

78. ———. Stress and disease. *Science,* 1955, **122**, 625-631.

79. SHERIF, M., & CANTRIL, H. The Psychology of Ego-Involvements. New York: Wiley, 1947.

80. SINNOTT, E. W., DUNN, L. C., & DOBZHANSKY, T. Principles of Genetics (4th ed.). New York: McGraw-Hill, 1950.

81. SOLOMON, F., & FISHMAN, J. R. Youth and social action: II. Action and identity formation in the first student sit-in demonstration. *J. Soc. Issues,* 1964, **20**, 36-45.

82. SOROKIN, P. A. Man and Society in Calamity. New York: Dutton, 1942.

83. ———. The American Sex Revolution. Boston: Porter Sargent, 1956.

84. SULLIVAN, H. S. A theory of interpersonal relations—The illusion of personal individuality. Paper presented to the Society on the Theory of the Personality, New York City, May 3, 1944. Reprinted in H. M. Ruitenbeek (Ed.), *Varieties of Personality Theory.* New York: Dutton, 1964.

85. WALTERS, C. E. Prediction of postnatal development from fetal activity. *Child Devel.,* 1965, **36**, 801-808.

86. WEBSTER, H. Changes in attitudes during college. *J. Educ. Psychol.,* 1958, **49**, 109-117.

87. WOLBERG, L. R. Hypnoanalysis. New York: Grune & Stratton, 1945.

88. ———. Hypnotherapy. In J. L. McCary (Ed.), *Six Approaches to Psychotherapy.* New York: Dryden, 1955.

89. WORCHEL, P., & BYRNE, D. (Eds.). Personality Change. New York: Wiley, 1964.

90. YARROW, L. J. Separation from parents during early childhood. In M. L. Hoffman and L. W. Hoffman (Eds.), *Review of Child Development Research* (*Vol. 1*). New York: Russell Sage Foundation, 1964.

91. YARROW, L. J., & YARROW, M. R. Personality continuity and change in the family context. In P. Worchel & D. Byrne (Eds.), *Personality Change.* New York: Wiley, 1964.

Department of Psychology
Louisiana State University
Baton Rouge, Louisiana 70803

[illegible]

[illegible]

[illegible] Amer., 1950, 66, [illegible]

[illegible]

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 193-198.

A STUDY OF FORM DISCRIMINATION IN THE OPOSSUM*[1]

The Department of Psychology, University of Georgia

W. T. JAMES AND JOHN MCFARLAND

A. INTRODUCTION

A number of experiments with the opossum have indicated the feasibility of using this animal for research. The opossum has little difficulty in solving problems that have been presented to other species (3, 4, 5, 6). One of the above studies dealt with a simple visual discrimination problem. In these tests the animal successfully discriminated between a black-white vertically striped card *versus* a black card (each card was 7½ × 11½ inches); and a black card with a small white triangle in the center *versus* a white card with a small black triangle in the center. Each triangle was equilateral with 1¼-inch sides. Although the animals probably responded to some aspect of white *versus* black in all the tests, the studies indicated that the opossum could be motivated to respond to visual signals. Recent studies by Friedman have also shown that the animal can be used in visual discrimination problems [see Friedman and Marshall (1)]. This report deals with further analysis of visual learning in the opossum in terms of object discrimination, and black-white discrimination of closely related signals.

B. METHOD

1. *Subjects*

Three animals, two females and one male reared by hand since old enough to leave the pouch, were used in the experiments. Since the animals had been handled while young, they were tame and responsive. One female was used in six tests, while the second female and the male were used in the last more complicated problem.

2. *Apparatus and Procedure*

During the training the animals were housed singly in a wire cage 2 × 2 × 4 feet high, with the bottom elevated 5 inches from the floor. A

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 2, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by the General Research Division of the University of Georgia.

sleeping quarter was located near the top of the cage with a wire ladder leading to its entrance. The animals remained asleep most of the daytime. However, they would leave the quarters at the regular feeding time, and it was at this time of the day that the tests were conducted. A large elevated plywood door was located on one side of the cage. This door could be raised to permit the animal to walk out on a platform constructed on that side of the cage. The discrimination apparatus was mounted on rollers to permit placement near or as far as 18 inches away from the door. The discrimination apparatus contained two levers, one right and one left, with a rectangular pan located between (see Figure 1). Food could be delivered to the pan by either lever. The device for delivering food was hidden behind the panel above the levers. The apparatus elicited the sound of a running motor as the lever was depressed, and rotating levers pushed meat down a chute to the pan. The discrimination objects were mounted vertically on pieces of

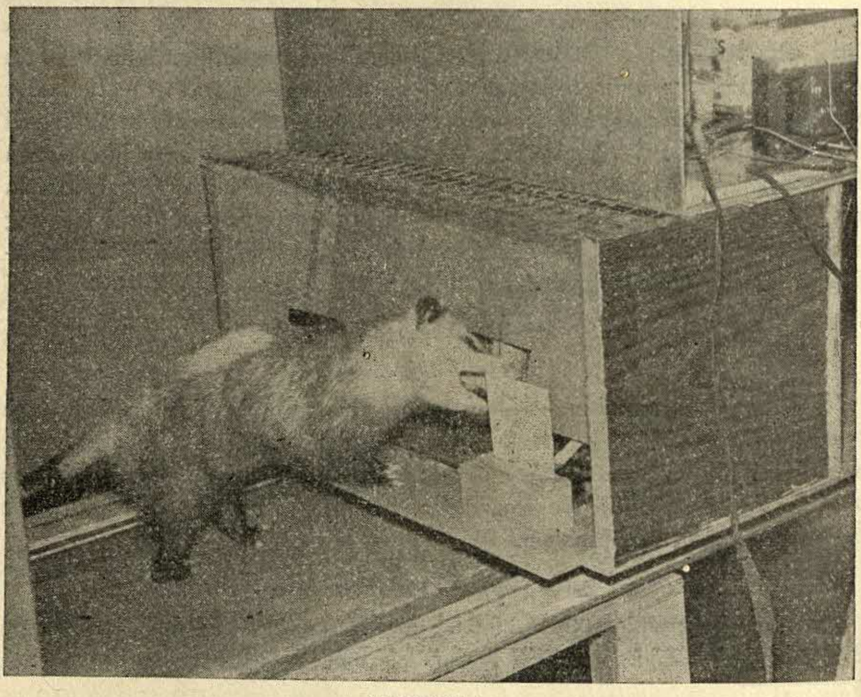

FIGURE 1
A Two-Choice Operant Discrimination Apparatus for the Opossum
The animal pushes against the positive white square in contrast to the black square on the left. As the white square is pushed, meat drops into the pan between the stimuli.

plywood 1/4 × 1/3 × 2 inches (see Figure 2), which could be placed in slots on the base of either lever in a matter of a few seconds. The background of the objects was painted a neutral gray. A 60-watt light was placed above the discrimination area.

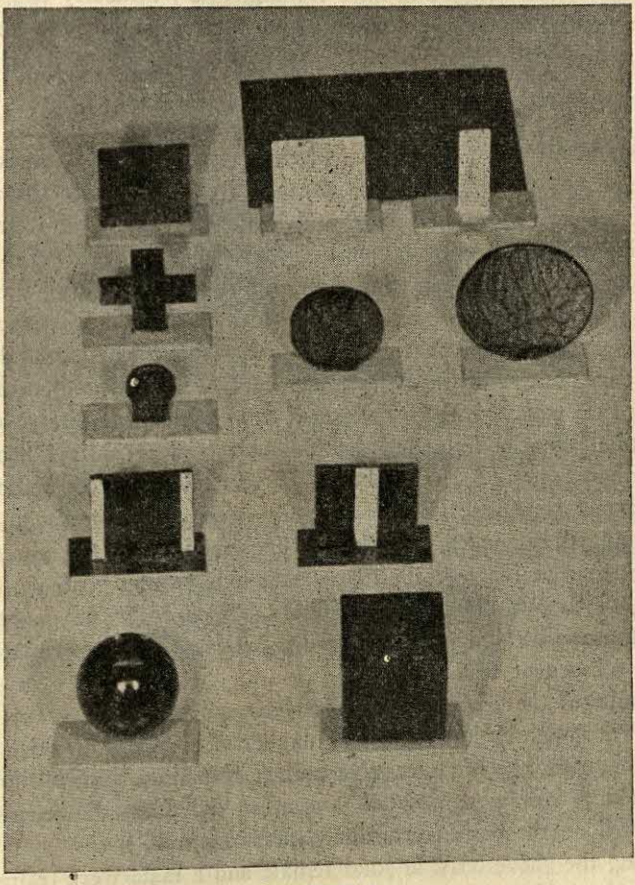

FIGURE 2
OBJECTS USED IN VISUAL DISCRIMINATION

The dimensions of the objects are as follows. Both squares in the first row are 3¼ × 3¼ inches. The small rectangle white object is 3¼ × 1½ inches; the black cross measures 3¼ inches each way, with tip of strips 1-inch across. The medium sized circle is 3¼ inches in diameter, the small black circle is 1⅝ inches in diameter, and the large circle is 4⅝ inches in diameter. The fourth row shows two squares the same size as those in the first row, but with ½-inch white strips of tape on the edge of one, and a white strip 1-inch wide in the center of the other. The last row shows a cube, sides 3 × 3 inches, which is contrasted with a ball of the same diameter.

Since the opossum is supposed to be mainly an anosmic animal, it is difficult to rule out all olfactory cues. In order to minimize this factor, however, pieces of meat used as reinforcement were placed at the top of each delivery tube, but only the one placed by the positive signal chute could be delivered. This procedure was assumed to have eliminated any olfactory cue that could be used in making the discrimination. The discriminanda were also washed at frequent intervals.

C. RESULTS

The experiment started with the mature female, using a black *versus* a white square. On the first test day the animal came down to the apparatus, already placed near the raised door, and learned to take food from the pan. After eating the meat, she was pushed back behind the lowered door for a few seconds before another trial. After 12 of these feedings, she returned to the cage on her own initiative. On the 13th trial, food was placed in one of the delivery tubes, and in the course of her movements around the pans she nosed the positive white object, and food was delivered automatically. This procedure was repeated on the next trial with the white object in the same position. On the following day the objects were shifted according to the Gellerman Series (2), with six out of 10 choices correct and three consecutive correct responses. For the next five series of 10 trials on successive days, the animal responded as follows: nine correct and five successive; 10 correct; nine successive correct; 10 correct; nine correct with six successive, for all of which $p < .03$ or above. There seemed to be no doubt that the two objects were discriminated, and results indicated the use of more complex objects in subsequent testing.

The results of the other tests on the above animal (Test group 1 to 6 in Table 1) are shown by indicating the number of blocks of 10 trials necessary to reach the criteria of six successive correct responses. Since 10 trial blocks were merely a convenient way of presenting the data, the various criteria are computed on the basis of continued performance.

Following the above tests, another female and a male were trained first on the black *versus* white squares. The female reached the discrimination criteria in five blocks, while the male required 17 blocks to reach criteria. The problem was made more complex at this point by placing a piece of 1-inch white tape on the outer edge of one black square, and a 1-inch white strip in the center of the other (see Figure 2). The female reached the proper criteria in three blocks, while the male obtained only a criteria of $p < .03$ in eight blocks.

The two objects were modified further with the strips on the positive

TABLE 1
NUMBER OF BLOCKS OF 10 TRIALS NECESSARY TO REACH CRITERIA OF SIX SUCCESSIVE CORRECT RESPONSES ($p < .01$) OR ABOVE, OR THE NUMBER OF BLOCKS GIVEN WITH FAILURE TO REACH CRITERIA[a]

Test group	Objects	Blocks to criteria, or failure of discrimination
1	Medium circle *vs.* cross	7 (three other blocks above this)
2	White square *vs.* white rectangle	Criteria not reached in eight blocks
3	Medium black disk *vs.* small disk	7 (criteria not reached)
4	Large black disk *vs.* small disk	2 (one block above this criteria)
5	Large black disk *vs.* medium disk	4 (one block above this criteria)
6	Cube *vs.* ball	6 (two blocks above this criteria)

[a] Details of the discriminated objects may be obtained from Figure 2.

object narrowed to ½ inch. There was immediate transfer on the part of the female, who reached the criteria in one block of trials. The male required two blocks of 10 trials.

The problem was further modified by moving the strips on the outer edge of the positive object in toward the center in two consecutive series of blocks of 10 trials each. The discrimination failed when the center strips were ¾-inch apart. The animal began to hesitate at this time, and the general discrimination dropped back to chance level. At this time the experiment was discontinued.

D. DISCUSSION

It is difficult to determine the exact basis on which the discrimination was made in the above tests. The separation of black *versus* white is apparently easy for the opossum. In most of the object discrimination tests there was no doubt that the animal had successfully distinguished between the two objects. This was apparent not only by the number of correct responses in the blocks of 10 as the test progressed, but also by the behavior of the animals. As the animal left the cage and approached the discriminanda vicarious trial and error behavior occurred in the early part of the trials. Later their movements indicated a direct choice. In the early part of some of the tests the opossum would many times push on the negative signal before shifting to the positive. As the experiment progressed, however, one push on the wrong object was enough to indicate failure and the animal would shift to the other immediately. In the correct discriminations the animal made the choice before reaching the objects, as indicated by the direction of his approach. As in all the other tests that have been done on the opossum, the present studies indicate better behavior than might have been expected of this animal.

E. SUMMARY

One female opossum, in an operant test situation, successfully discriminated between a black and white square with ease. Discrimination between a circle and cross required a large number of trials. The animal also had difficulty separating two small black disks that differed by 1⅝-inch diameter, but she could distinguish between black disks differing by 3 inches. Following the latter training, she successfully distinguished disks differing by only 1⅜ inches. A white square 3¼ × 3¼ inches *versus* a rectangle 3¼ × 1½ inches was an especially difficult problem, and the animal never did adequately distinguish between the two. A differential response between a ball and a cube offered no difficulty.

Another male and female were trained on the black *versus* white problem. After this, white strips of tape were placed on the edge of one black square and a 1-inch strip was placed in the center of another square of the same size. After the animals distinguished between these, the strips on the outer edge were reduced to ½ inch. Then the strips on the outer edge were moved inward toward the center by ¼ inch in two series of trials. The discrimination failed when the strips were ¾ inch apart in contrast to a 1 inch strip in the center of the second square.

REFERENCES

1. FRIEDMAN, H., & MARSHALL, D. A. Position reversal in the Virginia Opossum: Evidence for the acquisition set. *Quart. J. Exper. Psychol.*, 1965, 250-254.
2. GELLERMAN, L. W. Chance order of alternating stimuli in visual discrimination experiments. *J. Genet. Psychol.*, 1933, **42**, 207-208.
3. JAMES, W. T. The behavior of the opossum in the Guthrie-Horton puzzle box. *J. Genet. Psychol.*, 1955, **87**, 203-206.
4. ———. Conditioned responses in the opossum. *J. Genet. Psychol.*, 1958, **93**, 179-183.
5. ———. Behavior of the opossum in the Fink Arrow maze. *J. Genet. Psychol.*, 1959, **94**, 199-203.
6. ———. A study of visual discrimination in the opossum. *J. Genet. Psychol.*, 1960, **97**, 127-130.

The Department of Psychology
The University of Georgia
Athens, Georgia 30601

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 199-212.

THE EFFECTS OF ANXIETY, STRESS, AND TASK DIFFICULTY ON PERFORMANCE ON A VERBAL LEARNING TASK*[1]

Sylvania Electronic Systems

DONALD W. BROWN

A. INTRODUCTION

The purpose of this study was to assess the effects of stress on a verbal learning task as a function of the difficulty of the task and the anxiety level of the subject. Taylor (6) found that subjects scoring in the upper 12 percentiles on the Manifest Anxiety Scale gave a significantly higher percentage of conditioned responses than did subjects scoring in the lower nine percentiles, in an eyelid conditioning experiment. A similar increase in conditioned responses was found by Spence and Taylor (4).

Taylor and Spence (7) compared high and low anxious subjects on a verbal learning task. They found that the low anxiety group performed better (faster learning, fewer errors) than the high anxiety group. These results were interpreted as being the result of the opportunity for interference allowed in the multiple response situation: that is, the higher drive in the anxious subjects resulted in a decrement on this task because anxiety increases the strength of all responses in the organisms' repertoire, including the competing ones.

Montague (3) investigated the effect of anxiety upon experimental performance as a function of the relative number and strength of correct and incorrect tendencies elicited in the experimental situation. Three verbal learning tasks, made to vary in relative number of correct and incorrect tendencies elicited by the manipulation of intralist-similarity and association value of the nonsense syllables employed, were given to independent grouping of 20 anxious and 20 nonanxious subjects. Anxious subjects performed less well (number of correct anticipations) than nonanxious subjects on the difficult task with many incorrect tendencies, showed greater improvement of performance as the task became easier, and surpassed the nonanxious subjects on the task with the least number of incorrect tendencies.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 6, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was carried out while the author was attending Tufts University.

Deese, Lazarus, and Keenan (1) studied the relationships between experimentally induced stress and personality factors (score on a "neuroticism inventory" that correlates highly with the Taylor Scale of Manifest Anxiety) and a simple serial learning task. Ninety subjects practiced a list of 12 consonant nonsense syllables under three conditions of experimental stress: (*a*) shock administered for incorrect responses, (*b*) shock administered randomly for both correct and incorrect responses, and (*c*) no shock during training (control). Under the condition of shock for incorrect responses there was a marked difference favoring the anxious subjects. A smaller difference between these two groups was observed under the condition of random shock. This difference also favored the anxious group. Under the condition of no shock there was a very slight but consistent superiority of the anxious subjects. The authors hypothesize that the large difference in the "shock for incorrect response" situation is due to both a decrement in the performance of the low anxiety group and an increment in the performance of the high anxiety group. In the random shock situation, the difference is entirely due to a decrement in the low anxiety group.

Lazarus, Deese, and Hamilton (2) replicated the above study (1) with the exception that they used a more difficult serial learning task. This difference in task difficulty produced a very large difference in results. Not only were there no significant differences between the performance of the high and low anxiety groups, but the mean differences that were obtained were in the opposite direction to those found in the previous study (i.e., low anxiety subjects were superior to the high anxiety group).

In this study it is hypothesized that the presence of anxiety is a stressful state for the organism when it reaches high levels or when it interacts with a high level of task difficulty. Thus, while a high level of anxiety may be beneficial in simple learning tasks in which there are few competing response tendencies, as the number and strength of competing response tendencies increases subjects with a low anxiety level will perform better than subjects with a high level.

It was hypothesized that conditions of electric shock and task difficulty would combine to produce a stressful situation such that with low anxious subjects a relatively high level of stress would produce optimal performance, while with high anxious subjects optimal performance would result when the stress level was low (i.e., low anxious *S*s would be superior to high anxious *S*s on high stress conditions, while the opposite would be true under low stress).

B. METHOD

1. *Apparatus*

The Taylor Scale of Manifest Anxiety (6) was administered to 142 students in an introductory course in psychology and subjects were chosen randomly from males scoring in the upper and lower 20 per cent. These subjects were randomly assigned to one of three shock conditions (shock, threat of shock, no shock). The shock apparatus consisted of a 110 volt AC outlet, a step-down transformer, variable resistor, voltmeter, switch, and two silver electrodes. The two electrodes were strapped to the inside and outside surfaces of the subject's left forearm and a good contact was obtained through the use of electrode jelly.

The learning task consisted of three lists of nonsense syllables varying in similarity and association value. The lists were the same as those used by Montague (3). The lists were typewritten in capital letters on white tape and presented by means of a memory drum with an intersyllable time interval of three seconds and an intertrial interval of nine seconds.

2. *Procedure*

Eighteen subjects, nine high and nine low anxious were used. The high and low anxious subjects were further divided into three groups each, with a high and a low anxious group assigned to each of three shock conditions: (*a*) Shock, (*b*) Threat of Shock, (*c*) No Shock. Every subject received 25 trials on every list of nonsense syllables. The order in which the lists were learned was counterbalanced within each group of three subjects.

The subjects received the following instructions prior to the experiment:

> This is an experiment in verbal learning. A list of 12 nonsense syllables will appear one at a time in this slit (experimenter pointed to slit in memory drum in which the syllables appeared). There will be an interval of three seconds between each syllable and an interval of nine seconds between trials. Your task is to anticipate and spell the next syllable in the series when one appears. The syllables will always be presented in the same order. Make a response to each syllable whether you are sure or not.

The subjects in the "Threat of Shock" condition received these additional instructions: "For a certain number of wrong responses you will be given an electric shock. Failure to respond is counted as an error." The subjects in the "Shock" condition received the following additional instructions: "You will receive an electric shock for each wrong response. Failure to respond is counted as an error." The experimenter then answered any relevant questions the subject asked.

Prior to the beginning of the trials, the subjects in the "Shock" and "Threat of Shock" condition set the level of shock that they would receive in the experiment. The initial presentation of the shock was near threshold level and it was then increased until the subject reported that this was the strongest shock he would be willing to take.

This shock setting procedure was the only occasion at which subjects in the "Threat of Shock" group actually received a shock.

In the case of the "Shock" group they were allowed one trial through the list with no shock and thereafter were shocked for every incorrect response.

The subjects were not informed how many trials would be run on each list.

C. RESULTS

Table 1 summarizes the Analysis of Variance performed on the data obtained. From this table it can be seen that the only variable that had an effect significantly different from chance was that of List Difficulty. The effect of list difficulty is clearly shown in Figure 1, in which the mean number of correct responses are plotted as a function of trials for the three levels of list difficulty.

TABLE 1
SUMMARY OF ANALYSIS OF VARIANCE
(Performed on the data)

Source of variance	*df*	SS	*MSq*	*F*
Between subjects total	17	112.03		
Anxiety	1	10.49	10.49	1.68
Shock	2	20.52	10.26	1.64
Anxiety × shock	2	6.20	3.10	0.50
Error *b*	12	74.82	6.24	
Within subjects total	36	279.03		
List	2	216.03	108.02	53.74**
List × anxiety	2	2.26	1.13	0.56
List × shock	4	0.39	0.10	0.05
List × anxiety × shock	4	12.05	3.01	1.49
Error *w*	24	48.30	2.01	
Total	53	391.06		

** Significant at the .01 level.

The effect of anxiety on performance is shown in Figure 2. Inspection of this figure, plotting mean number of correct responses as a function of trials for the two levels of anxiety, shows that when the scores of the anxiety

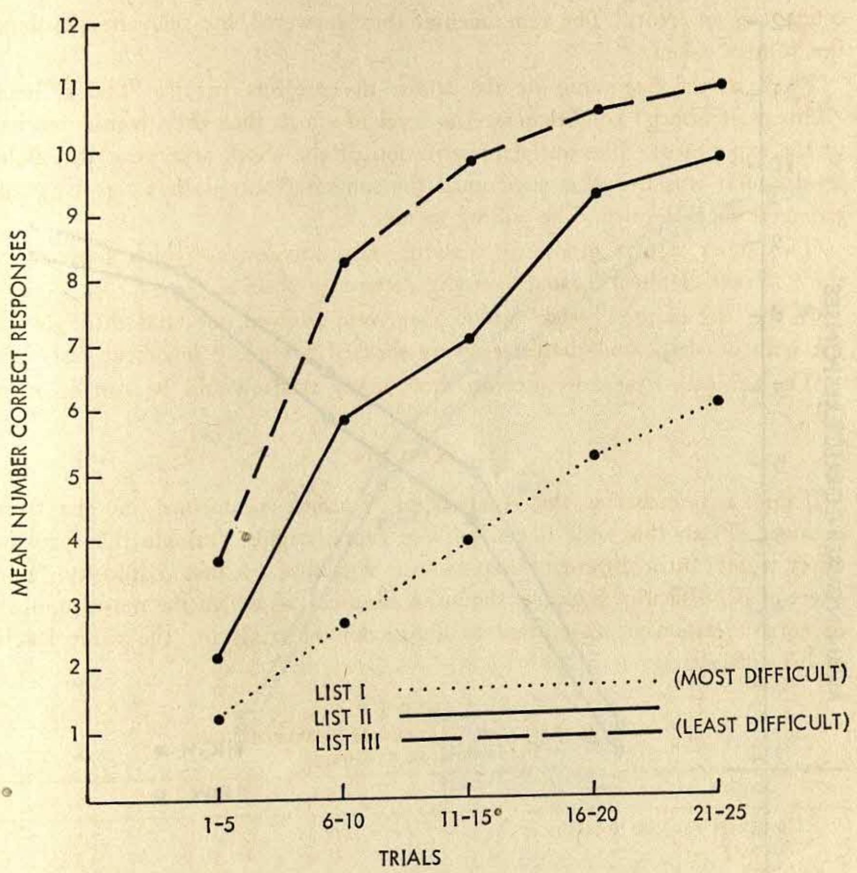

FIGURE 1
MEAN NUMBER CORRECT RESPONSES AS A FUNCTION OF TRIALS FOR THE THREE LEVELS OF LIST DIFFICULTY

groups are averaged over all experimental conditions there is a slight superiority on the part of the low anxious subjects.

Figure 3 is a plot of the mean number of correct responses as a function of trials for the shock, threat of shock, and no shock conditions. This figure indicates a slight superiority of the "Threat of Shock" condition, followed closely by the "No Shock" and at a lower level the "Shock" condition.

Figure 4 shows the mean number of correct responses as a function of trials for the two levels of anxiety and three levels of list difficulty employed. There is an indication that for the easy task, toward the end of training, the high

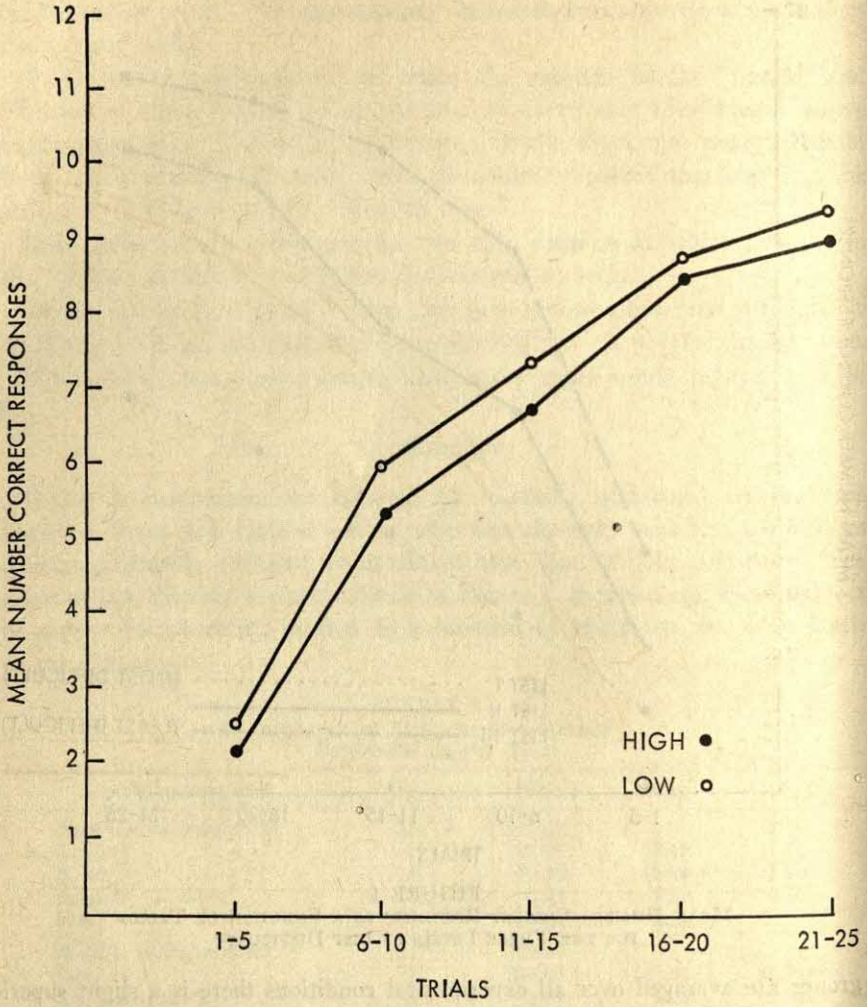

FIGURE 2
MEAN NUMBER CORRECT RESPONSES AS A FUNCTION OF TRIALS FOR THE HIGH AND LOW ANXIETY GROUPS

anxious group becomes equivalent to the low anxious. While the two are about equal on the intermediate list, the low anxious group was markedly superior, through all training trials, on the most difficult list.

The effects of anxiety level and shock condition on performance for Lists I, II, and III are given in Figures 5, 6, and 7, respectively. The relative positions

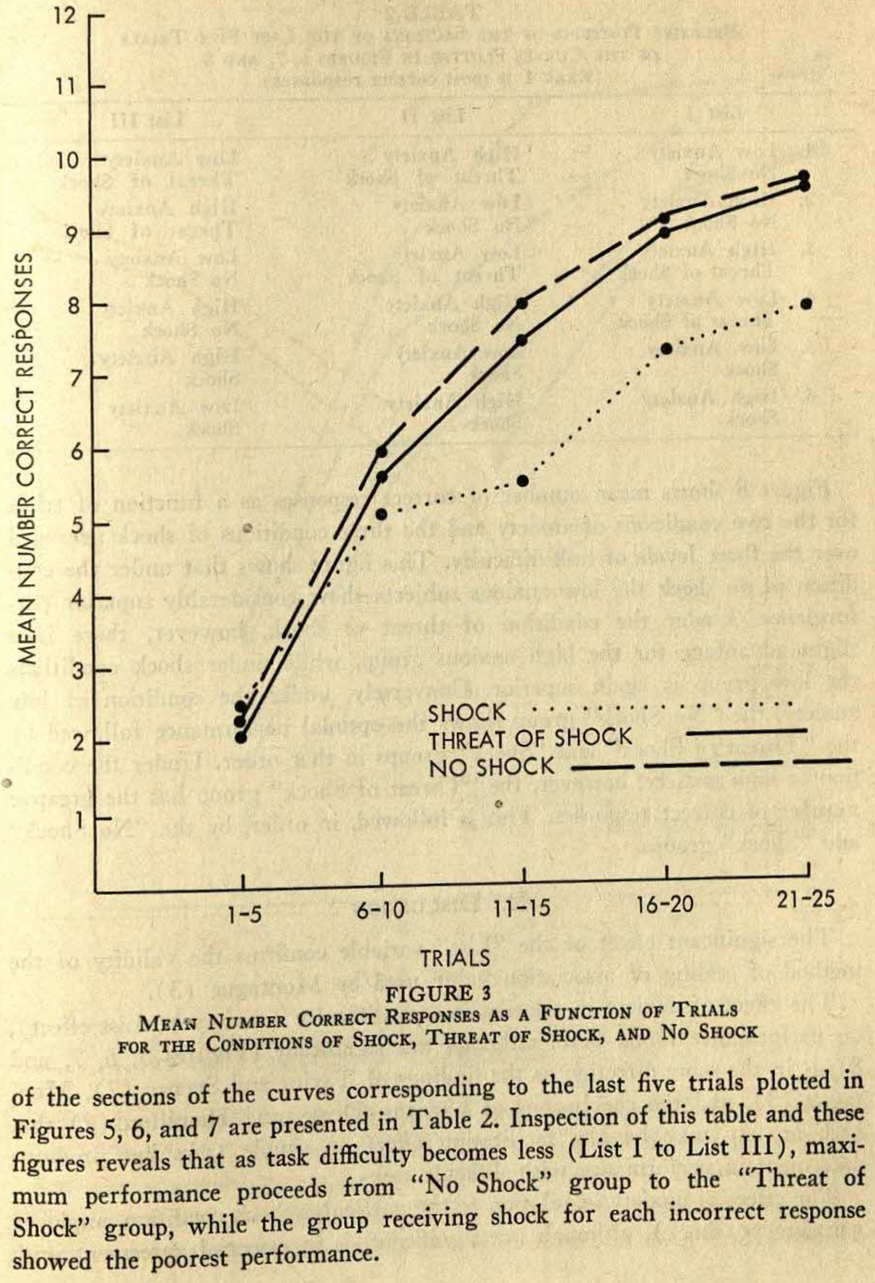

FIGURE 3
MEAN NUMBER CORRECT RESPONSES AS A FUNCTION OF TRIALS FOR THE CONDITIONS OF SHOCK, THREAT OF SHOCK, AND NO SHOCK

of the sections of the curves corresponding to the last five trials plotted in Figures 5, 6, and 7 are presented in Table 2. Inspection of this table and these figures reveals that as task difficulty becomes less (List I to List III), maximum performance proceeds from "No Shock" group to the "Threat of Shock" group, while the group receiving shock for each incorrect response showed the poorest performance.

TABLE 2
RELATIVE POSITIONS OF THE SECTIONS OF THE LAST FIVE TRIALS OF THE CURVES PLOTTED IN FIGURES 6, 7, AND 8
(Rank 1 is most correct responses)

	List I	List II	List III
1.	Low Anxiety No Shock	High Anxiety Threat of Shock	Low Anxiety Threat of Shock
2.	High Anxiety No Shock	Low Anxiety No Shock	High Anxiety Threat of Shock
3.	High Anxiety Threat of Shock	Low Anxiety Threat of Shock	Low Anxiety No Shock
4.	Low Anxiety Threat of Shock	High Anxiety No Shock	High Anxiety No Shock
5.	Low Anxiety Shock	Low Anxiety Shock	High Anxiety Shock
6.	High Anxiety Shock	High Anxiety Shock	Low Anxiety Shock

Figure 8 shows mean number of correct responses as a function of trials for the two conditions of anxiety and the three conditions of shock averaged over the three levels of task difficulty. This figure shows that under the condition of no shock the low anxious subjects show considerably superior performance. Under the condition of threat of shock, however, there is a slight advantage for the high anxious group, while under shock conditions the low group is again superior. Conversely, under the condition of low anxiety, the "No Shock" group shows the optimal performance followed by the "Threat of Shock" and "Shock" groups in that order. Under the condition of high anxiety, however, the "Threat of Shock" group has the greatest number of correct responses. This is followed, in order, by the "No Shock" and "Shock" groups.

D. DISCUSSION

The significant effect of the "List" variable confirms the validity of the method of scaling of association value used by Montague (3).

The effect of anxiety in relation to task difficulty (anxiety × List effect), on performance found here, although not significant (Figures 5, 6, 7, and 8), is in the same direction as the findings of Taylor and Spence (7), Montague (3), and the conclusions of Lazarus, Deese, and Hamilton (2) that high anxiety may have a facilitating effect on simple tasks but produces decrement in performance with difficult tasks.

The effect of anxiety in relation to the particular shock condition employed (anxiety × shock), although not significant, is in essential agreement with

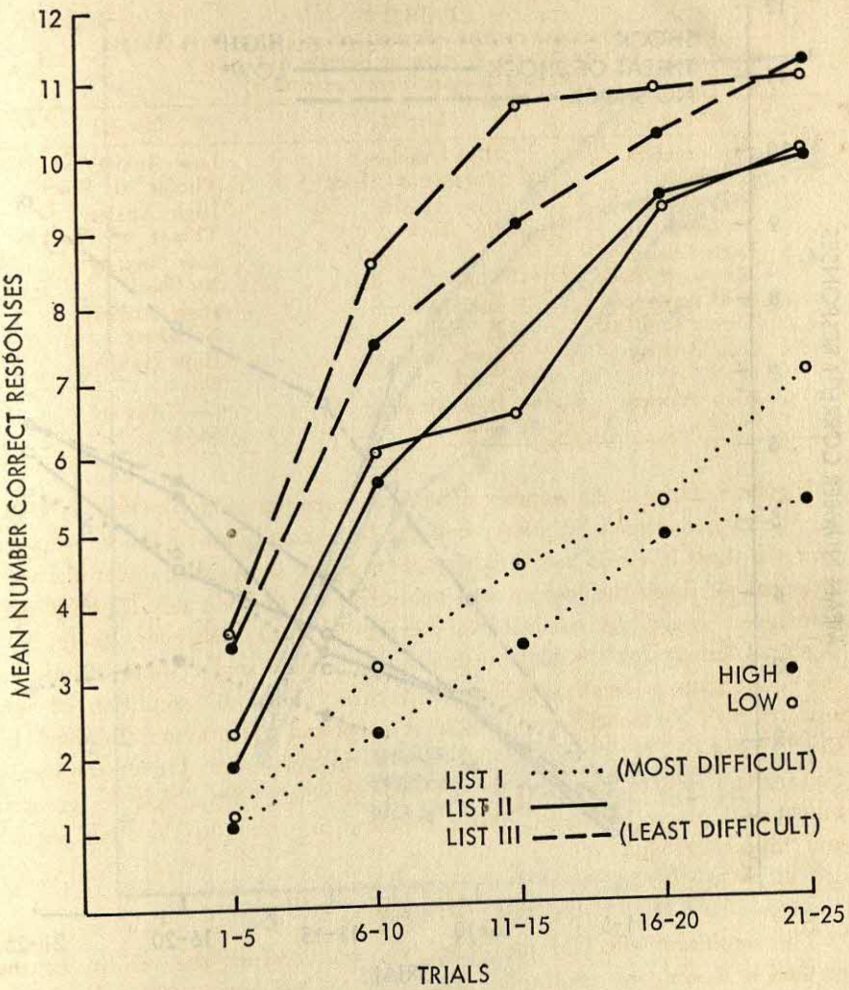

FIGURE 4
Mean Number Correct Responses as a Function of Trials for the Three Levels of List Difficulty and Two Levels of Anxiety

Spence, Farber, and Taylor (5). In that study three conditions of shock were used: "No Shock," "Threat of Shock," and a condition under which the subjects were shocked occasionally (at random) at the time of the ready signal. They found that the high anxiety groups that received shock performed at a higher level than those receiving no shock. In this study it was found that

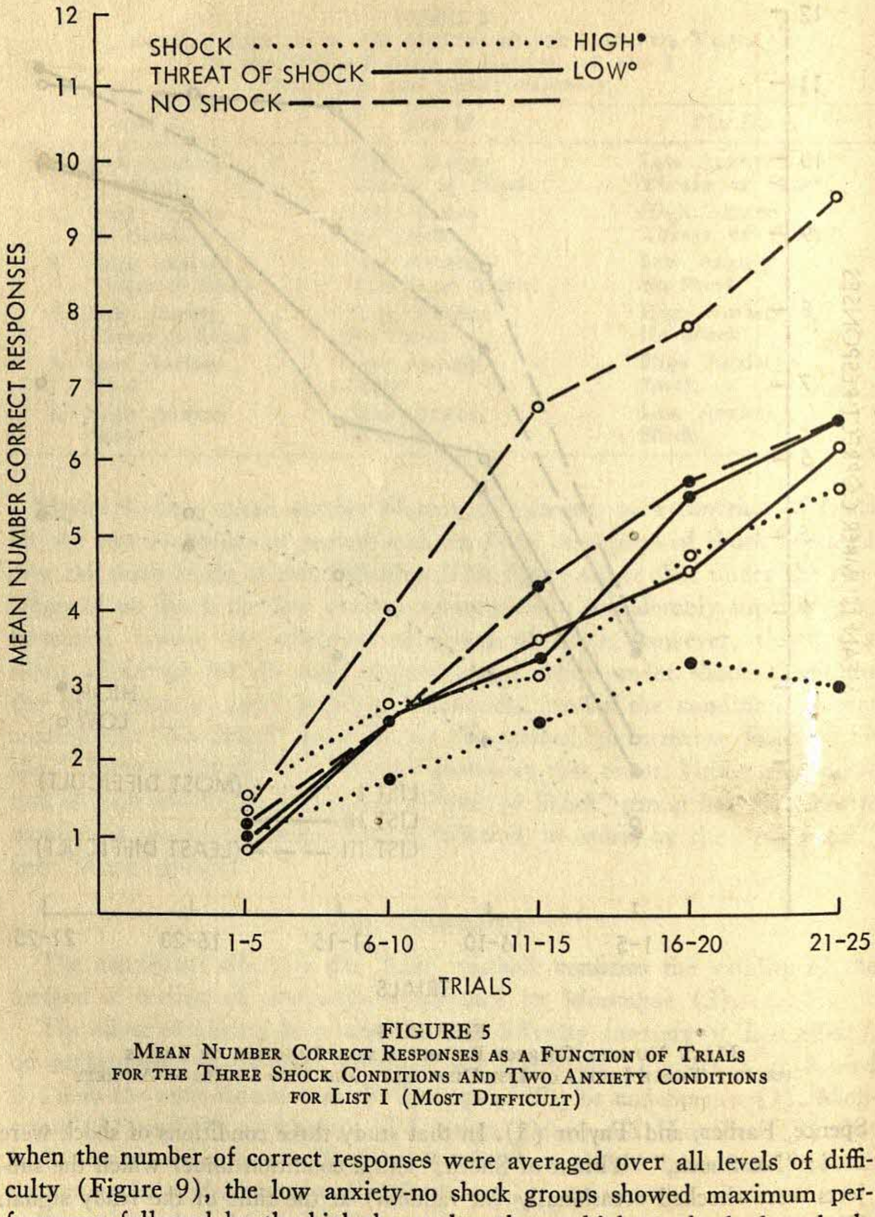

FIGURE 5
Mean Number Correct Responses as a Function of Trials for the Three Shock Conditions and Two Anxiety Conditions for List I (Most Difficult)

when the number of correct responses were averaged over all levels of difficulty (Figure 9), the low anxiety-no shock groups showed maximum performance followed by the high-threat, low-threat, high-no shock, low-shock, and high-shock groups. When, however, the effect of anxiety and shock of

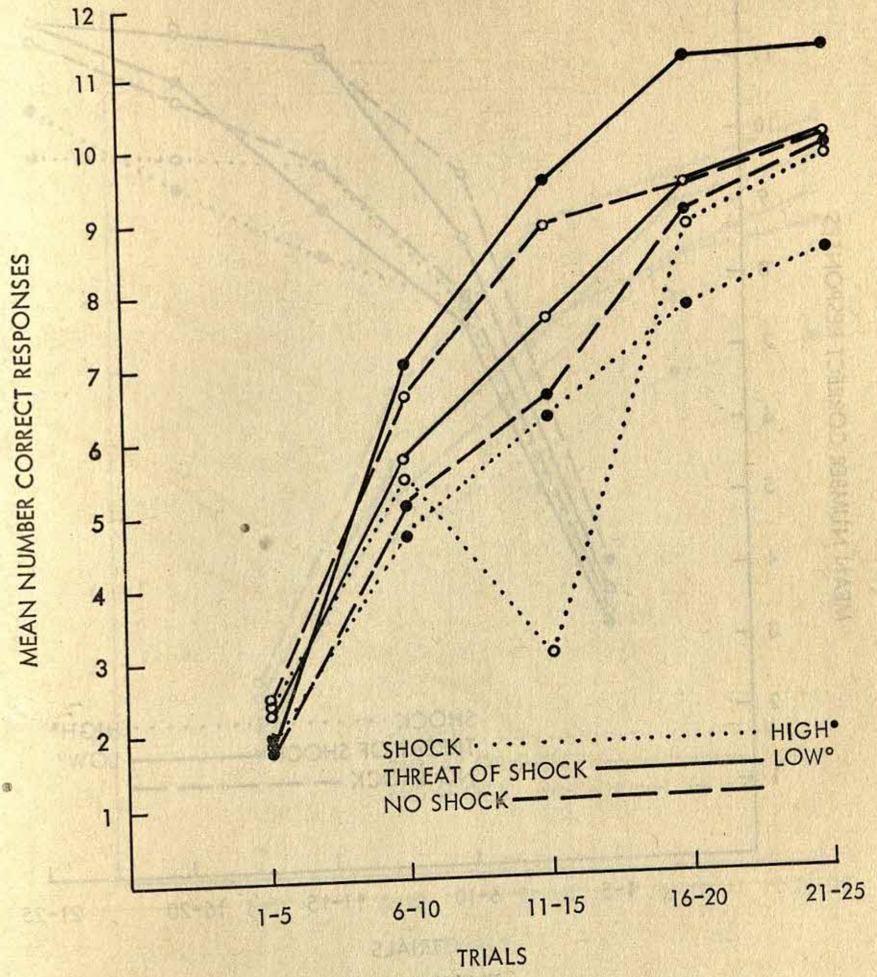

FIGURE 6
MEAN NUMBER CORRECT RESPONSES AS A FUNCTION OF TRIALS FOR THE THREE SHOCK CONDITIONS AND TWO ANXIETY CONDITIONS ON LIST II (INTERMEDIATE LEVEL OF DIFFICULTY)

performance are examined for the three levels of difficulty individually (Figures 6, 7, 8, and Table 2), it is seen that "Threat of Shock" increases in its ability to improve performance as task difficulty decreases. These findings may be interpreted as supporting those of Spence, Farber, and Taylor (5),

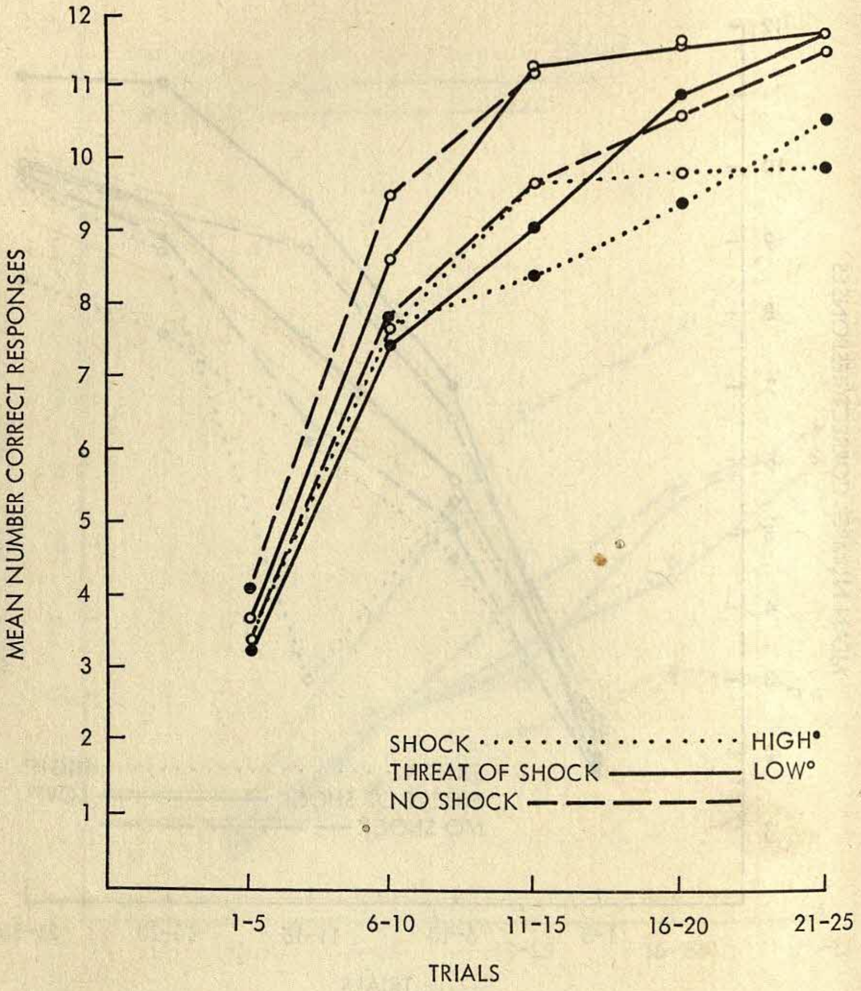

FIGURE 7
MEAN NUMBER CORRECT RESPONSES AS A FUNCTION OF TRIALS FOR THE THREE SHOCK CONDITIONS AND TWO ANXIETY CONDITIONS ON LIST III (LEAST DIFFICULT)

since in the Spence *et al.* study the task was one of eyelid conditioning, which is simpler than the easiest list used in this experiment (in terms of competing response tendencies). The degree of simplicity of the task can also account for the discrepancy between the present study and the Spence *et al.* study, with regard to the "Shock condition." Spence *et al.* (5) found

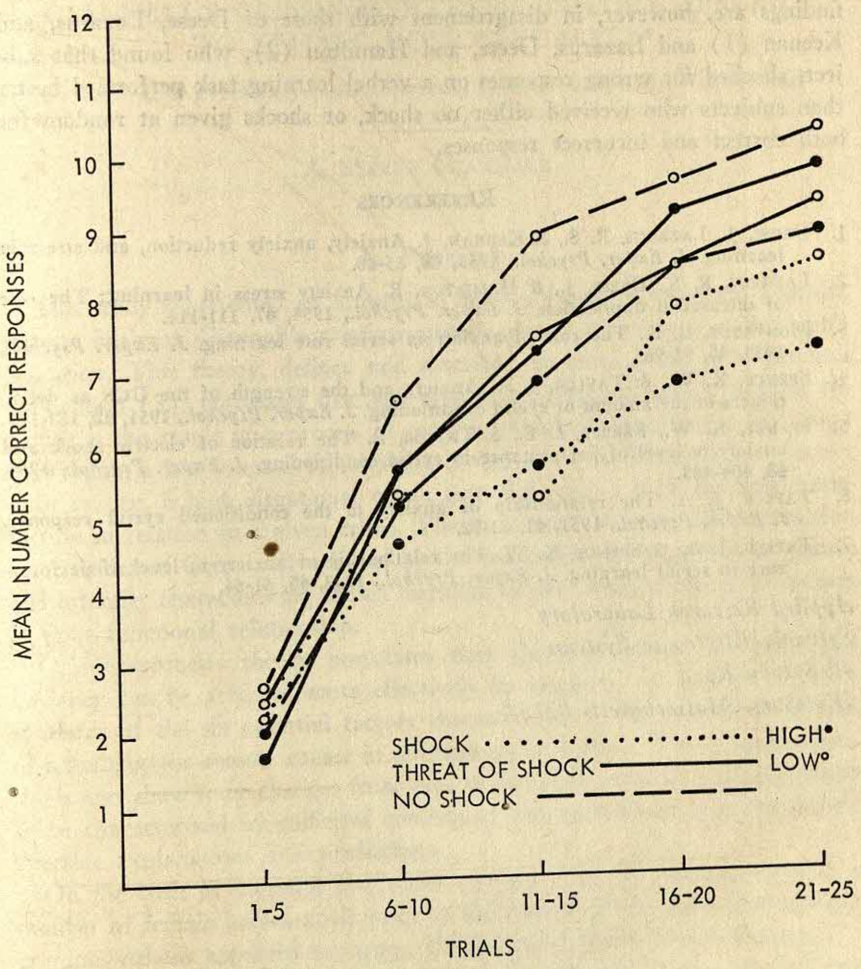

FIGURE 8
Mean Number Correct Responses as a Function of Trials for the Three Shock Conditions and Two Anxiety Averaged over the Three Lists

that shock administered occasionally with the warning signal facilitated performance on eyelid conditioning, while the present study found that shock for all erroneous responses produced the lowest level of performance observed (Figure 4). It appears that eyelid conditioning is sufficiently simpler than verbal learning to allow shock to produce near optimum performance. These findings are, however, in disagreement with those of Deese, Lazarus, and Keenan (1) and Lazarus, Deese, and Hamilton (2), who found that subjects shocked for wrong responses on a verbal learning task performed better than subjects who received either no shock, or shocks given at random for both correct and incorrect responses.

REFERENCES

1. DEESE, J., LAZARUS, R. S., & KEENAN, J. Anxiety, anxiety reduction, and stress in learning. *J. Exper. Psychol.*, 1953, **46**, 55-60.
2. LAZARUS, R. S., DEESE, J., & HAMILTON, R. Anxiety stress in learning: The role of intraserial duplication. *J. Exper. Psychol.*, 1954, **47**, 111-114.
3. MONTAGUE, E. K. The role of anxiety in serial rote learning. *J. Exper. Psychol.*, 1953, **45**, 91-96.
4. SPENCE, K. W., & TAYLOR, J. A. Anxiety and the strength of the UCS as determiners of the amount of eyelid conditioning. *J. Exper. Psychol.*, 1951, **42**, 183-188.
5. SPENCE, K. W., FARBER, I. E., & TAYLOR, E. The relation of electric shock and anxiety to level of performance in eyelid conditioning. *J. Exper. Psychol.*, 1954, **48**, 404-408.
6. TAYLOR, J. A. The relationship of anxiety to the conditioned eyelid response. *J. Exper. Psychol.*, 1951, **41**, 81-92.
7. TAYLOR, J. A., & SPENCE, K. W. The relationship of anxiety to level of performance in serial learning. *J. Exper. Psychol.*, 1952, **45**, 61-64.

Applied Research Laboratory
Sylvania Electronic Systems
40 Sylvan Road
Waltham, Massachusetts 02154

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 213-222.

GIANNELL'S CRIMINOSYNTHESIS THEORY APPLIED TO FEMALE HOMOSEXUALITY*

Department of Psychology, State University of New York, College at Potsdam

A. STEVEN GIANNELL

A. PROBLEM

This study represented an attempt to explain female homosexual behavior on the basis of Giannell's criminosynthesis theory of crime (or deviancy) causation. This theory, defined and described in more detail in a previous publication (3), postulates that any type of criminal or socially deviant behavior is the result of a process of cognitive synthesis involving a high need frustration, a low internal inhibition, a low external inhibition, a low contact with reality, a high situational crime potential, and a high potential satisfaction in relation to a given crime. The theory provides for the construction of an Index of Criminality yielded by the evaluation of the particular nature and intensity characterizing the six essential factors as perceived by the person in their functional relationship.

Criminosynthesis theory postulates that the explanation of crime and deviancy can be achieved more effectively by assessing the present cognitive synthesis of the six essential factors characteristic of the individual, instead of searching for remote causes in psychoanalytic fashion. This cognitive synthesis may show some changes from time to time, but every individual is likely to be characterized by sufficient consistency and persistency that can permit effective explanations and predictions.

On the basis of intensive interviews conducted by the investigator with a number of female homosexuals prior to the present study, their characteristic criminosynthesis appeared to express the following Index of Criminality.

Need frustration: a highly frustrated need to give and receive affection and love, as well as a high need for aggression, independence, and nonconformity.

Internal inhibitions: a complete lack of inhibitions regarding their homosexual status and practices; since they did not regard their behavior as either morally, socially, or criminally deviant, no internal forces could prevent them from engaging in such behavior.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 9, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

External inhibitions: no external forces prevented them from engaging in homosexual behavior, since in their environment no legal sanction could normally be imposed. Homosexual acts, however, are unlawful in Ohio and Michigan where the study was conducted, under the general provision of the sodomy statutes (4, 5). These statutes "generally apply to 'whoever,' 'any person,' 'every person,' etc., and consequently would seem to apply equally to female homosexuality and male homosexuality" (1). It is very difficult, however, to find legal evidence about these types of felonies, and "homosexuality, in and of itself, is not unlawful in any jurisdiction" (1).

Contact with reality: low, because the homosexual females seemed not to think or even care about their present or future life in a conventional family, occupation, or society; they appeared to be entirely preoccupied with their homosexual relationships, aside from which the reality of the conventional world did not exist or was perceived in an entirely distorted perspective.

Situational crime potential: high, since the opportunities to engage in homosexual relations were readily available in an environment where homosexual girls were plentiful, and could meet freely and publicly in a number of homosexual bars in the area.

Potential satisfaction: high, since their homosexual behavior provided the best possible satisfaction of their frustrated need for affection and nonconformity; it also provided a release of their aggressive needs in the chaotic, hyperactive, and often violent atmosphere of the homosexual bars without producing any appreciable loss in terms of material items or psychological conflicts.

In conclusion, the subjects' characteristic synthesis of their particular type and magnitude of need frustration, internal inhibition, external inhibition, contact with reality, situational crime potential, and potential satisfaction appeared to be highly conducive to homosexual behavior. A variation in the type and magnitude of any one among the six essential factors may have produced an entirely different synthesis and an entirely different behavior.

It would have been pertinent in this connection to formulate a differential-experimental design where homosexuals would be compared with nonhomosexuals in terms of all the six essential factors postulated by the theory. This was not possible, however, within the limits of the present investigation.

For this reason, only the factor referring to need frustration was selected for quantitative study in a differential-experimental type of research design, while the other factors became the subject of a qualitative analysis based on further interview data in a differential type of research design.[1]

---

[1] The terms "differential-experimental" design and "differential" design are used according to the definition given by Tinker and Russell (6).

In relation to the differential-experimental design, the following hypotheses were formulated. On the basis of the results obtained by the investigator from previous interviews, translated into Edwards terminology (2), it was hypothesized that when the Edwards Personal Preference Schedule is administered to a female homosexual sample and a comparison is made with the Edwards female college normative sample, the homosexuals are significantly higher than the normative sample on succorance, nurturance, autonomy, and aggression, as well as significantly lower than the normative sample on deference and heterosexuality. The Edwards female college normative sample was used instead of the female adult normative sample because the former was closer in age to the homosexual sample. The assumption was made that the normative sample did not contain homosexuals in any appreciable degree.

It was further hypothesized that when "Butch," "Femme," and "Neutral" subjects are compared with each other there are no significant differences among the three groups with regard to the needs mentioned. For the purpose of the present study, the term homosexual defines a female who engaged in homosexual sex practices at the time and for at least six months prior to the study. The term "Butch" generally defines a female who plays the part of the male in the homosexual relationship, while the term "Femme" refers to a homosexual who plays the part of the female; the term "Neutral" is used to designate a homosexual who plays the part of the male or of the female interchangeably.

In relation to the differential type of design, it was hypothesized that when a female homosexual sample is interviewed using criminosynthesis theory as a schedule, the homosexuals are characterized by a low internal inhibition, a low external inhibition, a low contact with reality, a high situational crime (deviancy) potential, and a high potential satisfaction with regard to homosexual behavior.

B. METHOD

The Edwards Personal Preference Schedule (2) was administered by the investigator to the female homosexual sample during the summer of 1964. This group constituted an incidental sample whereby all the subjects that made themselves available were tested. They were originally located in the "gay" (homosexual) bars in Toledo, Ohio and Detroit, Michigan. The subjects were residents of these cities or of surrounding areas, with the exception of two stripteasers who were traveling with a burlesque show, and three other girls who were residents of West Virginia. Most of the subjects were born and raised in the geographical area mentioned.

The mean age of the total sample was 22, with the "Butch," "Femme," and "Neutral" subgroups having a mean age of 22, 21, and 24, respectively. The mean educational level in terms of school grade completed was about 12, with quite a few subjects having from one to three years of college; some were college students at the time.

Occupationally, most subjects held jobs as waitresses, bar maids, beauticians, or office secretaries, either on a permanent or temporary basis. About 20 per cent were engaged in the fulltime practice of prostitution, while others supplemented their earnings with the proceeds of prostitution whenever needed.

About 5 per cent of the subjects had a criminal record, mainly for shoplifting, and a few were known to engage in criminal behavior at the time of the study, in terms of shoplifting or stealing money from customers while engaging in prostitution.

All subjects were white, female, American, with a proportion of ethnic and religious representation approximately reflecting the population of the area.

The Edwards Personal Preference Schedule was administered by the investigator in individual or group sessions either in his office, in the subjects' apartment, or in the homosexual bars when conditions were favorable. Most subjects were tested in the homosexual bars. All subjects were tested in the open: that is, no one was at the time confined in an institution.

The subjects' motivation regarding the test was excellent. This resulted from the fact that the test contains the heterosexuality scale in which the subjects were particularly interested, and from the fact that the subjects were, in general, intrigued to find out about their personality characteristics. A few days after the research was started, the subjects enthusiastically volunteered to take the test whenever the investigator announced a testing session in the homosexual bars.

The subjects were informed as to the characteristics of the test, as to the purpose of the study, and as to the availability of the results immediately or shortly after the administration. They were instructed to answer the test anonymously and to place a code name on the protocol along with the indication of their age, school grade completed, and "Butch," "Femme," or "Neutral" status.

While the results were given and explained to them, an interview was also conducted with the exception of a few cases. This interview was intended to gather information about their background, and about the other essential factors postulated by criminosynthesis theory.

The total number of subjects tested was 118, and the total number of valid tests 103. A test was considered valid when the consistency score was nine or more, following Edwards' (2) suggestion that tests having a consistency score of less than nine are of questionable validity. On the basis of this criterion, 10 tests were not valid, and the other five were incomplete protocols.

C. Results

In relation to the differential-experimental design, the statistical evaluation of the data was conducted by applying a series of *t*-tests to determine the significance of the difference of the total homosexual sample, as well as of the "Butch," "Femme," and "Neutral" subgroups in comparison with the female college normative sample of the Edwards Personal Preference Schedule on each of the 15 variables included in the test, plus the consistency score.

A series of *t*-tests was also applied to test the significance of the difference between three other subgroups divided on the basis of the heterosexuality score. This subdivision was done by placing the subjects having a heterosexuality raw score between zero and 10 in a low heterosexuality subgroup, those having a score between 11 and 18 in a medium heterosexuality subgroup, and those having a score between 19 and 28 in a high heterosexuality subgroup. These raw scores correspond to centile ranks of 0-24, 25-75, and 76-99 for the "Low," "Medium," and "High" subgroups respectively, in terms of the normative sample.

The results of all *t*-tests are reported in Tables 1 and 2. The *F* tests indicated that the variances were homogeneous in all comparisons.

The tables show that the first hypothesis (predicting significantly higher scores for the homosexual sample as compared with the normative sample on succorance, nurturance, autonomy, and aggression, as well as lower scores on deference and heterosexuality) was substantiated by the data with the exception of the succorance variable.

The second hypothesis (predicting no significant difference on the variables mentioned above among "Butch," "Femme," and "Neutral" subgroups) was also supported by the data. When the "Low," "Medium," and "High" heterosexuality subgroups were compared, some differences appeared, as shown in Table 2.

In relation to the differential design, the hypothesis (predicting low internal inhibitions, low external inhibitions, low contact with reality, high situational crime (deviancy) potential, and high potential satisfaction with regard to homosexual behavior) was confirmed by the qualitative data obtained in the interviews.

TABLE 1
VALUES OF *t* OF ALL HOMOSEXUALS AND OF SUBGROUPS *Versus* EDWARDS PERSONAL PREFERENCE SCHEDULE COLLEGE NORMATIVE SAMPLE

Need	All homosexuals	Butch	Femme	Neutral	Low	Medium	High
Achievement	*3.87***	*4.13***	.95	1.83	*4.21***	.22	1.08
Deference	5.36**	3.71**	3.44**	2.41*	3.15**	4.37**	*3.00***
Order	1.40	.82	1.10	.54	1.87	.49	1.07
Exhibition	1.04	*2.71***	1.57	.86	1.44	1.58	1.80
Autonomy	*3.24***	.54	*3.34***	*2.05**	*3.06***	1.42	.54
Affiliation	5.00**	4.24**	2.27*	2.54*	4.33**	.39	4.35**
Intraception	1.33	1.02	.47	.95	1.73	1.50	1.82
Succorance	.95	1.86	1.02	1.03	1.80	*1.99**	1.07
Dominance	3.10**	.78	3.42**	1.30	1.68	2.98**	1.49
Abasement	1.65	1.60	1.01	.23	*2.36***	.56	1.64
Nurturance	*2.61***	.39	*3.17***	1.08	1.63	*3.13***	.11
Change	2.57*	1.82	.92	2.02*	2.81**	.70	1.86
Endurance	*2.70***	*3.40***	.29	1.17	*3.75***	.84	.004
Heterosexuality	9.04**	7.01**	5.93**	3.07**	12.92**	1.02	*4.75***
Aggression	*8.18***	*5.44***	*5.53***	*3.48***	*6.79***	*3.72***	*3.34***
Consistency	.87	.55	.21	.92	.60	.06	1.32
N homosexuals	103	40	40	23	71	21	11
N normative	749	749	749	749	749	749	749

Note: Italics indicate significantly higher means for all homosexuals and for subgroups.

* Values of *t* significant at or beyond the 5 per cent level.

** Values of *t* significant at or beyond the 1 per cent level.

D. DISCUSSION

On the whole, the results confirmed the prediction made on the basis of the preliminary interview data. In terms of the Edwards Personal Preference Schedule data, evaluated on the basis of the comparison with the female college normative sample, the total homosexual sample appeared to be characterized by the following psychological needs.

They appeared to have a high need for nurturance, autonomy, and aggression, as well as a low need for deference and heterosexuality. On the need for succorance the homosexual sample was found not to differ from the normative sample, contrary to the prediction made by the investigator. In relation to the variables not considered in the author's prediction and on which significant differences were found, the homosexual sample appeared to have a high need for achievement and endurance, as well as a low need for dominance, affiliation, and change.

Analyzing the data in terms of the comparison between the "Butch," "Femme," and "Neutral" subgroups and the normative sample, each of the three subgroups appeared to be characterized by the following psychological needs.

TABLE 2
VALUES OF *t* OF HOMOSEXUAL SUBGROUPS COMPARED WITH EACH OTHER

Need	Butch *vs.* Femme	Butch *vs.* Neutral	Femme *vs.* Neutral	Low *vs.* medium	Low *vs.* high	Medium *vs.* high
Achievement	*2.42**	1.21	.93	*2.02**	.61	.91
Deference	.20	.29	.14	*2.37**	1.59	.15
Order	.20	.07	.25	.49	1.81	1.06
Exhibition	*3.30***	.83	1.38	1.91	.98	2.40*
Autonomy	1.80	1.30	.40	.23	.65	.32
Affiliation	1.46	.46	.58	1.75	*2.24**	*3.19***
Intraception	.43	.14	.56	*2.47**	1.17	3.28**
Succorance	1.77	.31	1.26	2.31*	.26	*2.08**
Dominance	*1.98**	.59	1.06	1.90	.79	.51
Abasement	.38	.74	.38	.66	*2.08**	1.68
Nurturance	1.84	.61	.92	1.79	.64	1.73
Change	.73	.64	1.16	2.38*	.82	*2.21**
Endurance	1.85	1.09	.68	*2.27**	1.23	.54
Heterosexuality	.78	1.58	.99	10.90**	16.98**	10.06**
Aggression	.11	.56	.59	.07	.59	.51
Consistency	.25	.40	.67	.37	1.07	1.08
N top of headings	40	40	40	71	71	21
N bottom of headings	40	23	23	21	11	11

Note: Italics indicate significantly higher means for subgroups at top of headings.
* Values of *t* significant at or beyond the 5 per cent level.
** Values of *t* significant at or beyond the 1 per cent level.

The "Butch" subgroup appeared to have a high need for achievement, exhibition, endurance, and aggression, as well as a low need for deference, affiliation, and heterosexuality; the "Femme" subgroup a high need for autonomy, nurturance, and aggression, as well as a low need for deference, dominance, affiliation, and heterosexuality; and the "Neutral" subgroup a high need for autonomy and aggression, as well as a low need for deference, affiliation, change, and heterosexuality.

Upon further analysis of the data in terms of the comparison between the "Low," "Medium," and "High" heterosexuality subgroups and the normative sample, each of the three groups appeared to be characterized by the following psychological needs.

The "Low" subgroup appeared to have a high need for achievement, autonomy, abasement, endurance, and aggression, as well as a low need for deference, affiliation, change, and heterosexuality. The "Medium" subgroup appeared to have a high need for succorance, nurturance, and aggression, as well as a low need for deference and dominance; and the "High" subgroup a high need for deference, heterosexuality, and aggression, as well as a low need for affiliation.

In terms of the other variables included in criminosynthesis theory—i.e., internal inhibition, external inhibition, contact with reality, situational crime potential, and potential satisfaction—the qualitative data obtained from interviewing the homosexual sample was in line with the data obtained from interviewing a number of homosexuals prior to this study.

In addition to a high need frustration, the homosexual sample was in fact characterized by low internal inhibitions, low external inhibitions, low contact with reality, a high situational crime (deviancy) potential, and a high potential satisfaction in relation to homosexuality. Further differential-experimental research is suggested in order to find quantitative data from a comparable homosexual sample in relation to all the variables included in criminosynthesis theory.

Whenever the correct criminosynthesis can be identified and a valid Index of Criminality constructed in relation to a given crime or deviancy category, the criminal or deviant tendencies of a subject can be assessed beforehand on the assumption that the psychological characteristics of the individual are sufficiently persistent and consistent.

On the basis of this assessment, a plan directed to the rehabilitation of the offender can be outlined. Although one cannot change the nature of the six essential factors that criminosynthesis postulates to be characteristic of each individual, a change in behavior can be achieved by a cognitive restructuring of such factors.

Regarding the homosexual females, in particular, the investigator observed, for example, that the need for nurturance appeared to have been frustrated for a long time in the "Butch" subjects because they never had any relationships with the opposite sex mainly due to their unattractive physical appearance. In the "Femme" subjects, on the other hand, this need appeared to have been frustrated because for a long time these subjects, generally very attractive, had been treated by men simply as sex objects, without any regard for their psychological needs. Both types of subjects developed an aversion to men for different reasons, and channeled their need satisfaction in the direction of the same sex.

A rechanneling in the direction of the opposite sex could be obtained by changing the girls' attitudes toward men in a process of cognitive restructuring whereby the "Butch" would be influenced to understand that physical unattractiveness is not to be equated with total rejection from men, and the "Femme" subjects to realize that a number of unpleasant experiences with men should not be generalized to a negative evaluation of all men.

The same process of cognitive restructuring could be achieved in relation to the other essential factors postulated by criminosynthesis theory. This could be achieved either in individual or group sessions, whereby the subjects would arrive at an assessment of their particular type of criminosynthesis, and devise a way to restructure one or more of the six essential factors to produce a noncriminogenic type of synthesis.

E. SUMMARY

This study attempted to explain female homosexual behavior on the basis of Giannell's criminosynthesis theory of crime (or deviancy) causation. This theory postulates that any crime or deviant behavior is the result of a high need frustration, a low internal inhibition, a low external inhibition, a low contact with reality, a high situational crime potential, and a high potential satisfaction in relation to a given crime. The theory provides for the construction of an Index of Criminality yielded by the evaluation of the nature and intensity of the six essential factors in their functional relationship.

On the basis of intensive interviews conducted by the investigator with a number of female homosexuals prior to the present study, their Index of Criminality (deviancy) appeared to be characterized by a highly frustrated need for affection, love, aggression, independence, and nonconformity; it appeared also to be characterized by a low internal inhibition, a low external inhibition, a low contact with reality, a high situational crime (deviancy) potential, and a high potential satisfaction in relation to homosexual behavior.

Since, within the limits of the present study, a differential-experimental design comparing homosexuals with nonhomosexuals on all of the six essential factors was not possible, only the factor of need frustration was selected for a differential-experimental quantitative study, while the other essential factors became the subject of a qualitative analysis in a differential type of research design.

In relation to the differential-experimental quantitative study, it was hypothesized that when the Edwards Personal Preference Schedule is administered to a female homosexual sample, and a comparison is made with the Edwards female college normative sample, the homosexuals are significantly higher on succorance, nurturance, autonomy, and aggression, and significantly lower on deference and heterosexuality. It was further hypothesized that when the "Butch," "Femme," and "Neutral" subjects are compared with each other, there are no significant differences among the three subgroups with regard to the needs mentioned.

In relation to the differential type of design, it was hypothesized that when a female homosexual sample is interviewed using criminosynthesis theory as a schedule, the homosexuals are characterized by a low internal inhibition, a low external inhibition, a low contact with reality, a high situational crime (deviancy) potential, and a high potential satisfaction with regard to homosexual behavior.

A total of 103 subjects were tested and interviewed by the investigator in the open. The hypotheses formulated in relation to the differential-experimental design were confirmed by the data, with the exception of the need for succorance included in the first hypothesis. The hypothesis formulated in relation to the differential design was also confirmed by the qualitative data obtained from the interviews. It was concluded that criminosynthesis theory appears to be useful in the explanation of crime and deviant behavior.

REFERENCES

1. CANTOR, D. J. Deviation and the criminal law. *J. Crim. Law, Criminol., & Police Sci.,* 1964, **55**, 441-453.
2. EDWARDS, A. L. Edwards Personal Preference Schedule Manual. New York: Psychol. Corp., 1959.
3. GIANNELL, A. S. Psychological needs characteristic of four criminal-offender groups. *J. Soc. Psychol.,* 1966, **69**, 55-72.
4. Michigan Statutes Annotated. Chicago: Callaghan, 1962.
5. Ohio Revised Code Annotated. Cleveland: Banks-Baldwin, 1958.
6. TINKER, M. A., & RUSSELL, W. A. Introduction to Methods in Experimental Psychology. New York: Appleton-Century-Crofts, 1958.

Department of Psychology
State University of New York
College at Potsdam, N.Y. 13676

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 223-226.

THE EFFECT OF DNA ON OPERANT CONDITIONING*[1]

Department of Psychiatry, McGill University

C. BEAULIEU AND L. SOLYOM

A. INTRODUCTION

The aim of this study is to test the hypothesis that desoxyribonucleic acid (DNA) facilitates learning of a bar-pressing response for water reinforcement.

Studies with ribonucleic acid (RNA) have shown that it has a facilitating effect on learning and retention (1, 2). This was most clearly shown using operant conditioning methods (6).

In contrast with the rapid turnover rate of RNA, the DNA content of the cellular nuclei is remarkably constant; the rate of DNA turnover is not changed by excitation and inhibition of the central nervous system, electroshock, administration of phenobarbital sodium amytal, hibernation, or by changes of the oxygen supply of the brain [reported by Gaito (5)]. However, Gaito (4, 5) postulates that during a behavioral event the DNA becomes active, making RNA and consequently increasing protein synthesis. He claims that in addition to the functionally inactive DNA, a functionally active DNA exists, as shown by more recent experiments.

Since DNA is similar in structure to RNA, differing only in the sugar component of the molecular chain, it can be hypothesized that it also would facilitate learning (4). An additional experimental group was used that received RNA in addition to DNA in order to see if the combined use of the two would lead to further improvement in learning functions. The control group received a saline solution.

B. METHOD

1. *Design*

A unifactorial, three-group design was used in this study. The animals were assigned at random to each of the three groups.

2. *Subjects*

Originally 10 animals were assigned to each of the three groups. All were white female rats of the Wistar strain, approximately four months old at

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 9, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study was supported by the Dominion Provincial Mental Health Grant No. 604-5-76.

the onset of the experiment. Due to technical difficulties, a total of 10 animals had to be rejected from the analysis, leaving seven animals in the saline group, eight in the DNA group, and five in the DNA + RNA group.

3. *Preparation of DNA and RNA*

Fish sperm DNA was obtained from Schwartz Bioresearch Inc., and yeast RNA from Pabst Laboratories. Ten per cent neutral solutions were prepared by the addition of 16 M NaOH to pH 7.

4. *Injections*

160 mg/kg of DNA was administered intraperitoneally to the DNA groups: the DNA + RNA group received 80 mg/kg of each substance, and the saline group received an equivalent volume, .5 cc, of saline. All injections were intraperitoneal and were given at the end of each day, five days a week, for a period of approximately three weeks prior to onset of the experiment and throughout the entire course of the experiment.

5. *Procedure*

All animals were first subjected to a master training technique developed by Corson (3), which essentially consists of placing a naive animal with an animal that has already been trained in the operant conditioning situation. This method has been shown to lead to faster acquisition of the bar pressing response than is obtained with the more conventional shaping methods. Following master training, the animal was to be tested alone the following day. If the animal met a criterion of five or more responses during a one-half-hour period, he was then placed on continuous reinforcement schedule. Failure to meet this criterion either during testing or during continuous reinforcement training was followed by retraining on the following day. All animals were exposed to four consecutive days of continuous reinforcement training. After this, they were exposed to a two-and-one-half-minute variable interval schedule for five consecutive days. Finally, this was followed by 14 consecutive days of extinction.

C. RESULTS

No significant difference between the three groups could be found during master training. A chi square test comparing the number of animals in each group who required more than two days of master training with those requiring only one day suggested that the control animals required more training than did the other two groups. However, this finding was far from being significant (chi square = 1.36, *df* = 2).

An analysis of variance on the mean number of bar presses for each group during the four days of continuous reinforcement revealed no group effect ($F < 1$) nor any significant group by trial interaction ($F = 1.55$, $df = 6/51$, $p < .10$). However, a significant trials effect was found ($F = 4.88$, $df = 3/51$, $p < 0.01$), which means that training was effective in increasing the mean number of bar presses in all three groups (see Figure 1).

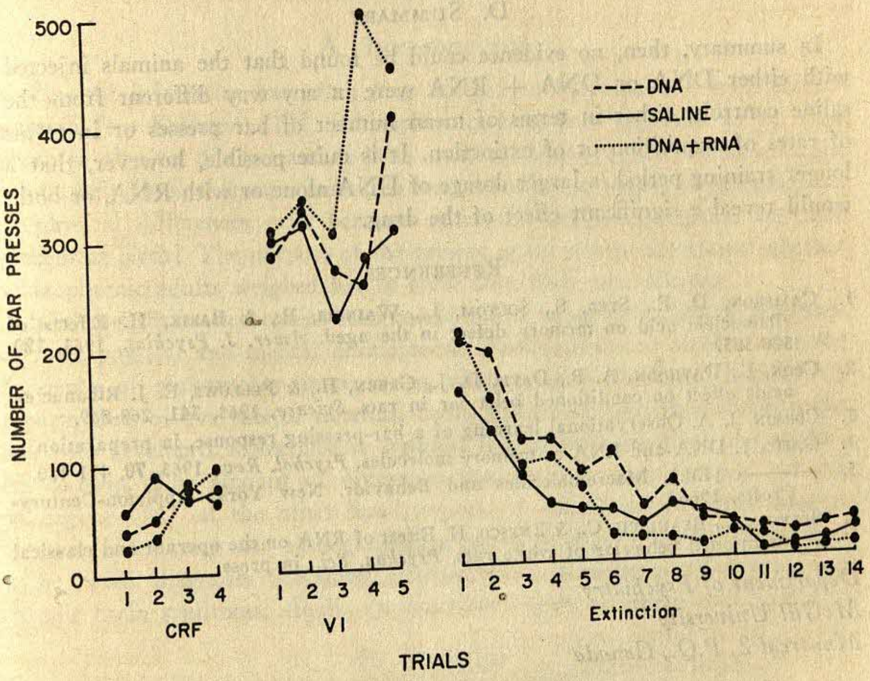

FIGURE 1
ANALYSIS OF VARIANCE ON THE MEAN NUMBER OF BAR PRESSES FOR EACH OF THE THREE GROUPS

Analysis of variance on the mean number of bar presses for each of the three groups during the five days of variable interval training indicates that the three groups were not different from each other ($F = 1.66$, $df = 2/17$). The interaction was also nonsignificant ($F = 1.52$, $df = 8/76$). Surprisingly enough, a significant day-effect was found ($F = 4.48$, $df = 4/76$, $p < .01$). Closer analysis of this data was then carried out by means of Duncan's Multiple Range Test to find the exact nature of this effect. It was found that the only two days that were significantly different from each other were numbers

3 and 5. No differences could be found between any other pair of days. This suggests the possibility of an inverted relationship between amount of bar pressing and number of days of training. An analysis of variance performed on the extinction data revealed, as expected, a highly significant trial effect ($F = 46.87$, $df = 13/221$, $p < 0.001$). However, no significant group effect or interaction effect could be found.

D. SUMMARY

In summary, then, no evidence could be found that the animals injected with either DNA or DNA + RNA were in any way different from the saline controls, either in terms of mean number of bar presses or in terms of rates of acquisition or of extinction. It is quite possible, however, that a longer training period, a larger dosage of DNA alone or with RNA, or both, would reveal a significant effect of the drug.

REFERENCES

1. CAMERON, D. E., SVED, S., SOLYOM, L., WAINRIB, B., & BARIK, H. Effects of ribonucleic acid on memory defect in the aged. *Amer. J. Psychiat.*, 1963, **120**, 320-325.
2. COOK, L., DAVIDSON, A. B., DAVIS, D. J., GREEN, H., & FELLOWS, E. J. Ribonucleic acid: effect on conditioned behavior in rats. *Science*, 1963, **141**, 268-269.
3. CORSON, J. A. Observational learning of a bar-pressing response, in preparation.
4. GAITO, J. DNA and RNA as memory molecules. *Psychol. Rev.*, 1963, **70**, 471-480.
5. ——— (Ed.). Macromolecules and Behavior. New York: Appleton-Century-Crofts, 1966.
6. SOLYOM, L., BEAULIEU, C., & ENESCO, H. Effect of RNA on the operant and classical conditioned behavior of white rats, *Psychon. Sci.*, in press.

Department of Psychiatry
McGill University
Montreal 2, P.Q., Canada

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 227-231.

COMPARATIVE BIRTH WEIGHTS OF SCHIZOPHRENICS AND THEIR SIBLINGS*[1]

Psychology Department, Western Reserve University

ELLEN A. LANE AND GEORGE W. ALBEE

A. INTRODUCTION

Physical inferiority has often been alleged to be associated with schizophrenia [see Kallman (1)]. Schizophrenics have been observed to be more prone to physical diseases and to have a shorter life expectancy than controls, while their physical stature has often been described as slight, or ectomorphic. If physical differences exist between schizophrenics and controls, are they present at birth? The purpose of the present study was to determine whether schizophrenic adults weighed less at birth than their own siblings.

Birth weight has sometimes been considered to be an indication of prognosis of later physical and mental development. Pasamanick (2) reports that the incidence of abnormality increases as birth weight decreases. Pollin (4) reports that of five sets of identical twins discordant for schizophrenia, the one who developed schizophrenia weighed less at birth in all instances than his co-twin. The amount of difference was not reported. Pollack and his colleagues (3), on the other hand, report no differences in birth weights between index cases and siblings from a group of approximately 40 subjects with mixed diagnoses including schizophrenia, personality disorders, and chronic brain syndrome. Again, no exact data were presented.

B. METHOD

Birth certificates of 52 schizophrenic adults and their 115 siblings were located in the Bureau of Vital Statistics at Cleveland City Hall. The names and birthdates of schizophrenics had been obtained from several state and federal hospitals in the Cleveland area. The number of subjects in the sample was limited by the requirement that they must have attended the Cleveland public schools during childhood, for it was only through school attendance

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 12, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The research reported in this paper was supported by Grant M-5186 from the National Institute of Mental Health, United States Public Health Service, Department of Health, Education and Welfare.

records that names and birthdates of siblings could be obtained. Subjects and siblings also had to have been born in Cleveland. Fifteen subjects were female, 37 were male (reflecting the use of two large VA hospitals). Fifty-nine siblings were females, 56 were males. Birth certificates of subjects and siblings born since 1945 include a record of birth weight. Birth certificates before 1945 specified the hospital in Cleveland where birth occurred. Birth weights for subjects born before 1945 were obtained from the original records of four large general hospitals. The birth weight of each schizophrenic-to-be was compared to the average birth weight of his own siblings.

C. RESULTS

Table 1 shows that the birth weights of 52 schizophrenics averaged approximately six ounces less than the average birth weights of their own siblings. The average birth weight of the schizophrenics was almost seven pounds, while the average birth weight of the average of their siblings was approximately seven and one-half pounds. This difference is statistically significant at better than the .01 level. The birth weights of almost 70 per cent of the schizophrenic group were lower than the average birth weights of their siblings, and 61 per cent of the subjects weighed the very least in their family at birth no matter what the number of siblings, as shown in Table 2. The probability that the latter distribution could happen by chance is below the .01 on a X^2 test.

TABLE 1
MEAN BIRTH WEIGHTS IN OUNCES OF 52 ADULT SCHIZOPHRENICS AND THEIR SIBLINGS ($p = .01$)

Statistic	Schizophrenics	Average siblings
M	111.36	117.20
σ	18.79	17.57

Because it was found that five schizophrenic subjects of this group were premature at birth, while only five of the larger group of siblings were premature, the data were reanalyzed with the exclusion of all babies weighing less than five pounds, to insure that the differences found were not heavily dependent upon the influence of these premature subjects. Table 3 shows that even when the birth weights of full-term infants who later became schizophrenic are compared to the average birth weights of their full-term nonaffected siblings, the difference between the groups remains statistically significant at the .01 level. The average difference between the birth weights of schizophrenics and their siblings was still approximately six ounces.

TABLE 2

OBSERVED (O) AND EXPECTED (E) FREQUENCIES OF 52 SCHIZOPHRENICS-TO-BE WEIGHING THE LEAST, SOMEWHERE IN THE MIDDLE, OR THE MOST AMONG THEIR OWN SIBLINGS ($p = .01$)

N children in family		Lowest birth weight	Middle birth weight	Highest birth weight
2	O	17	0	5
	E	11	0	11
3	O	9	3	3
	E	5	5	5
4	O	4	4	1
	E	2.25	4.5	2.25
5	O	0	1	0
	E	.20	.60	.20
6	O	2	1	0
	E	.50	2	.50
7	O	0	1	0
	E	.143	.714	.143
10	O	0	1	0
	E	.10	.80	.10
Total	O	32.000	11.000	9.000
	E	19.193	13.614	19.193

As there might have been some question that a large number of first-born schizophrenics had resulted in the lower average birth weights of this group, the data excluding prematures were further recalculated with the exclusion of all first-born children. Although the average difference between the birth weights of schizophrenics and their siblings was now approximately four and one-half ounces, as shown in Table 4, this difference was consistent enough to reach the .05 level of significance on a one-tailed test.

TABLE 3

MEAN BIRTH WEIGHTS IN OUNCES OF 47 ADULT SCHIZOPHRENICS AND SIBLINGS, WEIGHING AT LEAST FIVE POUNDS ($p = .01$)

Statistic	Schizophrenics	Average siblings
M	114.92	121.29
σ	15.92	15.24

TABLE 4

MEAN BIRTH WEIGHTS OF 25 FULL-TERM NONFIRST-BORN ADULT SCHIZOPHRENICS AND THEIR SIBLINGS ($p = .05$)

Statistic	Schizophrenics	Average siblings
M	116.48	121.05
σ	16.60	14.46

A final analysis, not presented here, of the birth weights of the 37 males and 15 female subjects showed that sex differences were not relevant to the results; schizophrenics of both sexes were surpassed in birth weights by their siblings by approximately the same amount.

D. Discussion

This study lends support to the view that future schizophrenics start life at some disadvantage compared with their siblings. If birth weights were entirely unrelated to schizophrenia, no consistent differences would have been found between the birth weights of schizophrenics and those of their siblings. Instead, approximately 70 per cent of the schizophrenic group weighed less at birth than the average of the other children in their family, and the average differences between the groups was six ounces. As it is quite likely that some of the siblings used as normal controls also became schizophrenic in adulthood, the true differences in birth weight might well be even larger than those found.

On the other hand, it is difficult to see just how such physical "inferiority" could be directly related to schizophrenia. Indeed, these subjects, with an average birth weight of near seven pounds, would not be considered low except in comparison to their own siblings.

Biologists frequently use birth weight as a measure of the influence of environmental variables operating during pregnancy. The environment of the future schizophrenic individual begins with conception, and the process of birth is merely a transition from one environment to another. Ricks and Nameche (5) found that traumatic pregnancies differentiated mothers whose infants later became schizophrenic adults from control mothers. Traumatic pregnancies included "crisis levels of mental disorder in the mother during pregnancy or immediately following the birth of a child, extreme disruption in family life at the time of the pregnancy, serious physical illness in the mothers, and illegitimacy." It is reasonable to hypothesize that these conditions influence the birth weight of the infant who subsequently becomes a schizophrenic, and that this particular infant arrives in the world weighing less than the average of his siblings.

This hypothesis would also suggest that because this particular birth is associated with a period of tension that exceeds that of other pregnancies in the same mother, she is less well organized to cope with the endless demands for attention and care of this particular child, and that the mother-child relationship in this instance is inferior to the relationship with the other children. Some corroboration for this hypothesis was also reported by Ricks and Nameche (5), who found that mothers with infants that later became schizophrenic could be differentiated further from control mothers by the frequency of "an obsessive fear that the child would not survive delivery, inability to identify the child as one's own, perception of the child as either unrealistically perfect and ideal . . . or as unrealistically bad . . . ." Pollin (4) similarly reported that mothers of identical twins later discordant for schizophrenia perceived the future schizophrenic as being in greater peril, needing more protection, attention, and involvement than his co-twin, but this perception was realistic in the majority of cases. As a consequence, however, a pattern of greater dependency and decreased competence developed in the index twin.

Although it is speculation whether the reported differences in pregnancy experiences and maternal attitudes between mothers of schizophrenics and mothers of nonschizophrenics would also be found between the pregnancies of women who give birth to both schizophrenic and nonschizophrenic offspring (a study that would be of interest), such a hypothesis illustrates one possible alternate to a genetic one to account for the finding of lower birth weights for schizophrenics than for their own siblings.

E. Summary

Birth weights of 52 schizophrenic adults were found to average approximately six ounces less than the average birth weights of their own siblings. Significant differences in birth weights between schizophrenics and their siblings are not due to factors of prematurity or birth order. An alternate explanation to a genetic one is offered.

References

1. Kallman, F. J. Heredity in Health and Mental Disorder. New York: Norton, 1953.
2. Pasamanick, B., & Knobloch, H. Complication of pregnancy. In G. Caplan (Ed.), *Prevention of Mental Disorders in Children.* New York: Basic Books, 1961.
3. Pollack, M., Woerner, M. G., Goodman, W., & Greenberg, I. M. Childhood developmental patterns of hospitalized adult schizophrenic and non-schizophrenic patients and their siblings. *Amer. J. Orthopsychiat.*, 1966, **36**, 510-517.
4. Pollin, W., Stabenan, J. R., & Tupin, J. Family studies with identical twins discordant for schizophrenia. Paper presented at the annual meeting of the American Psychiatric Association, Los Angeles, California, May 7, 1964.
5. Ricks, D. F., & Nameche, G. Symbiosis, sacrifice, and schizophrenia. Paper presented at the annual meeting of the National Association for Mental Health, New York, 1965.

Psychology Department
Western Reserve University
Cleveland, Ohio 44106

[illegible] found [illegible] than [illegible] [illegible] the child as deaf [illegible] perfect and ideal . . . or as exceptionally bad . . . [illegible] reported that mothers of [illegible] [illegible] greater dependency and decreased competency [illegible]

[illegible] it is [illegible] whether the reported differences in pregnancy [illegible] and abnormal attitudes between mothers of [illegible] and [illegible] of [illegible] would also be found between the pregnancies [illegible] (a [illegible] that would be of interest), and a [illegible] out to account for the [illegible] of [illegible] birth weights [illegible] for their own siblings.

E. SUMMARY

Birth weights of 32 schizophrenic adults were found to [illegible] than the [illegible] birth weights of their [illegible] siblings. [illegible] between [illegible] and their siblings are due to [illegible] of prematurity on birth order. An alternate explanation [illegible] is offered.

REFERENCES

1. [illegible] Heredity in Health and Mental Disorder. New York: Norton, [illegible].
2. [illegible] B., & [illegible] H. Complications of pregnancy. In G. Caplan (Ed.), Prevention of Mental Disorders in Children. New York: Basic Books, [illegible].
3. [illegible]
4. [illegible] Paper presented at the [illegible] meeting of the American Psychiatric Association [illegible]
5. [illegible] Paper presented at the annual meeting of the National Association for Mental Health, New York, [illegible].

Psychiatry Department
Western Reserve University
Cleveland, Ohio 44106

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 233-239.

BRAIN BLOOD-SHIFT THEORY: INTERSENSORY RIVALRY, TIME ERROR, AND CLOSURE ILLUSIONS*

Department of Psychology, DePaul University

CONRAD CHYATTE

A. INTRODUCTION

E. D. Adrian (1) showed that electrical activity induced by stimulating a point on the exposed brain cortex of a rabbit tends to activate surrounding tissue. He observed these spreading electrical fields over areas of brain surface as large as a square centimeter. Such experimental data support efforts by W. Köhler (6) to account for closure, time-error effects, and other Gestalt phenomena in humans.

When two briefly separated images of the same line segment are flashed on a screen, the observer may report a positive time error: i.e., he may say that he judges the first image to have been longer (in extensity) than the second. According to Köhler, each image activates its isomorphic pattern on the visual cortex. By the time the second image is flashed, the initial pattern has enlarged through spreading excitation. Köhler held that, superimposed on this larger area of activity, the second field of brain excitation is experienced as the smaller of the two; should the second pattern spread to exceed the waning initial field, a negative time error is to be expected.

Closure effects also are explainable in terms of spreading cortical excitation. For example, look directly at the intersection of white strips in the very center of Figure 1. The white-strip intersections at other points will show grey closure patches. By Köhler's theory, these patches fall at points in the visual field that are isomorphic with cortical *loci* where spreading excitation overlaps maximally. Note, however, that the directly fixated intersection retains its white appearance; closure fails to occur there. Apparently excitation fails to spread efficiently at the topologically equivalent cortical *locus* for the point of visual fixation. These differences are predictable through brain blood-shift theory (3).

B. BRAIN BLOOD-SHIFT THEORY

Direct experimental evidence demonstrates that selective, localized displacement of blood in the brain occurs in response to stimulation through different

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 12, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

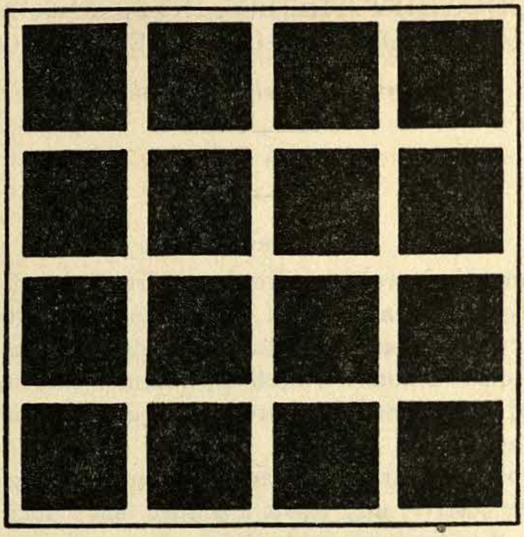

FIGURE 1
CLOSURE ILLUSION
Closure illusion appears at all white intersections but the one fixated (see text for brain blood-shift interpretation).

sensory modalities. For example, C. F. Schmidt (12) found increased blood flow to the cat's occipital cortex in response to visual stimuli. A. V. Lorenzo *et al.* (7) showed that the concentration of blood-borne radioactive sulfate increased appropriately in the visual or auditory cortex of the cat, depending on whether stimulation was by light or by sound waves. It is reported by H. E. Himwich (4) that "such an adaptive vascular response is observed when the retina is stimulated by light and both the lateral geniculate body and the visual cortex show signs of increased blood flow" (4, p. 104).

In developing brain blood-shift theory, the writer has stated that brain neurons compete for blood to nourish heightened metabolic demand under excitation; that they shift blood from their surround by activating adjacent perivascular neuroglia (e.g., astrocytes). First posited by S. Ramon y Cajal (10), this intrinsic cerebral mechanism is held to depend on neuroglial contractions that locally dilate brain capillaries. Some mechanism of this sort does function, and has been verified experimentally by D. H. Ingvar (5). With cats under reticular-system arousal, he used canalization to observe directly increases in cortical circulation that were independent of changes in blood pressure.

C. BLOOD-SHIFT GRADIENT AND INTERSENSORY RIVALRY

The writer has based a number of now verified predictions from brain blood-shift theory on the notion that cortical tissue closest to the excited *locus* should be deprived of the most blood, with a falling gradient of vascular deprivation at more distant points in the brain. The geometric extent of the gradient should widen as the intensity of stimulation or excitation increases. Thus, intersensory rivalry may be predicted on the basis of blood shift.

For example, the somatosensory (postrolandic) gyrus projects tactual *loci* that fall at varying distances from the auditory (temporal) cortex. According to T. Rasmussen and W. Penfield (11), there is a sequence of somatosensory *loci* in the following order: ankle (farthest from the auditory brain), knee, finger, cheek, upper lip, lower lip, teeth and gums, and inside of throat (closest to the auditory area). The writer therefore predicted that a relatively intense stimulus (electric shock) would act to draw increasing amounts of blood from the auditory cortex and thereby produce an increasing hierarchy of auditory thresholds as the point of cutaneous shock was moved along the sequence from ankle to inner throat. On the basis of this suggestion, two of the writer's undergraduate students[1] performed the following experiment as their research project for the course in Experimental Psychology.

Ten of their fellow students reported on their auditory thresholds (as measured with a Whipple audiometer) under four different conditions: (*a*) without shock, (*b*) with shock from an Electrodyne stimulator to the right leg (about 7.5 cm from the medial malleolus of the ankle on the inner calf), (*c*) with shock to the palmar side of the tip of the right index finger, and (*d*) with shock to the cheek (about 2.5 cm below the external canthus of the left eye). Even with the relatively low voltages employed for subject comfort (cheek, three volts; finger, 25 volts; leg, 20 volts), verification of the gradient in tactual-auditory rivalry as predicted from blood-shift theory was found. Arithmetic mean auditory threshold with shock to the cheek was significantly higher ($p < .01$) than the mean for the without-shock condition, and significantly higher ($p < .01$) than the mean for the finger-shock condition. Mean difference between finger and leg was not statistically significant.

On the strength of this pilot study, one of the writer's graduate students[2] performed a more detailed study to investigate the predicted gradient (8). He used a sample of 26 subjects, measuring auditory thresholds with a Beltone model 10C audiometer and administering shocks with a Harvard

---

[1] Kathleen Elizabeth Coulson Mele and Bonnie Lee Rosemarie Anderson.

[2] Jesse A. Macro.

Stimulator model 342. For a tone of 500 cps, this student found these mean thresholds in decibels (shock intensity shown in parentheses): 26.89, no shock; 27.57, ankle (90 volts); 27.85, knee (90 volts); 27.80, finger (120 volts); 28.26, cheek (25 volts); and 28.45, lip (20 volts). All thresholds at these voltages were significantly higher ($p < .01$) than that for the no-shock condition. The rank-difference correlation coefficient between this hierarchy of thresholds and that predicted on the basis ot a blood-shift gradient is +.94.

Since some auditory pathways decussate, each ear is represented by cortical tissue on both brain hemispheres. The writer therefore advised Robert M. Shannon (13) that bilateral shock (e.g., left and right sides of the lower lip shocked together) should shift blood more efficiently from the auditory cortex on each hemisphere, producing a greater threshold elevation than does unilateral shock. Shannon used a tone of 125 cps, worked in an anechoic room, and varied voltages in keeping with the maximum tolerance of each one of his 15 subjects. His results are shown in Table 1.

TABLE 1
MEAN AUDITORY THRESHOLD (DECIBELS) WITH SIMULTANEOUS ELECTRIC SHOCK

Type of shock	Cheek	Upper lip	Lower lip	Upper gum	Tongue
Unilateral shock (left side)	10.17	11.01	11.65	12.44	13.57
Bilateral shock (both sides)	11.33	12.15	12.44	13.33	14.60

Shannon reported a mean threshold of 7.59 decibels for the no-shock condition. The two threshold hierarchies represented by his data are in full agreement with the predictions based on the hypothesized blood-shift gradient. Considering the number of permutations possible for two such arrays of six means combined (including the no-shock mean), there is only one chance in 16,588,800 that such a set of data could have arisen at random.

This sort of intersensory rivalry is not limited to auditory-tactual function. For example, S. Novak (9) has reported that cutaneous electric shock increased the tendency of people to perceive two briefly separated flashes of light as one.

Competition for blood between disparate brain areas thus can be understood to generate visual-tactual (occipital-postrolandic) or auditory-tactual (temporal-postrolandic) rivalry. Also, blood shift within the same brain area may be expected to produce intrasensory rivalry (e.g., intraoccipital for visual function and intratemporal for auditory activities). The common difficulty people have in attending simultaneously to two or more conversations is offered here as an example of intratemporal rivalry. In the same way, the ostensibly anomalous or contradictory presence and absence of visual closure at different points in Figure 1 may be attributed to intraoccipital competition for blood.

D. Blood-Shift Gradient and Closure

Maximal blood shift should occur toward the small part of the occipital cortex that is involved in seeing the point of visual fixation in Figure 1. Visual neurons in closest proximity to that cortical *locus* should suffer the greatest deprivation of metabolites, therefore tend to be unresponsive, locally block the spread of excitation, and thus abort the viewer's experience of closure. Under ordinary illumination (relatively low stimulus intensity), the effective geometric limits of the blood-shift gradient should be much narrower than those to be expected under more intense local excitation. The writer therefore suggests that closure under such low-intensity stimulation is most probable for brain *loci* that are far enough away from the site of maximal blood shift to retain nourishment sufficient for supporting the spread of excitation. On the basis of such a blood-shift gradient, closure is most likely to be seen at *loci* in the visual field that are peripherally removed from the fixation point (and which indeed appear as grey patches); and least probable close to the point itself (which, in fact, remains white).

Smaller, closer black squares than appear in Figure 1 may provide peripheral white-strip intersections to stimulate brain neurons that fall within the effective blood-shift gradient. The peripheral intersections then also should fail to induce closure. Or the limits of the gradient may be expanded by sharply increasing the intensity of illumination, likewise making the peripheral intersections appear white. The same result seems feasible with substantial intersensory rivalry, perhaps using comparatively high-voltage cutaneous shock. Step back a yard or so from Figure 1 to put the whole image within the gradient and the patches will fade. Under brighter light they will disappear at a lesser distance.

A phenomenon that seems related to visual time error and closure has been reported by G. Stanley (14). He found that people who induced an afterimage, and then viewed it for a time with their eyes closed, tended to see it as having grown larger when their eyes were reopened. Blood-shift theory apparently accounts for this, since, when the eyes are closed to reduce metabolic demand in the visual system, blood should become more readily available throughout such structures as the retina and visual brain, promote spreading electrical excitation, and produce an expanded afterimage.

E. CONCLUSIONS

Brain blood-shift theory promises to provide remarkably powerful predictive and explanatory concepts. It seems to embrace psychological phenomena that classically have been approached with apparently unrelated hypotheses that tend to discourage synthesis. This report has presented theoretical and experimental arguments that act to unify such manifestations as the interference of cutaneous electric shock with visual and auditory function under the same theoretical mantle as Gestalt formation (e.g., closure and time error). Empirical data offered verify the prediction from blood-shift theory that human auditory acuity will deteriorate most when cutaneous shock excites sectors of the somatosensory brain that are geometrically closest to the temporal (auditory) lobe.

In earlier papers (2, 3) the blood-shift theory has been invoked to predict or interpret the transition from unconscious to conscious activities during hypnosis or in waking psychotherapy; pain control in audioanalgesia; hallucinatory experiences during sensory deprivation and in dreaming; perceptual deterioration associated with atherosclerosis in aging; and electroencephalographic manifestations associated with perceiving, introversion-extraversion, and anxiety.

Combined with better-known neurological concepts, localized blood shift seems to explain what the writer calls the alien-word phenomenon, in which the observer who stares at a printed word begins to see it as strangely unfamiliar. The psychoanalytic mechanism of repression also may be interpreted in terms of local blood shift. On the other hand, the notion of massive blood shift (to the brain as a whole) may be used to predict that verbal learning and retention will be facilitated when multiple modes of acquisition are employed (e.g., simultaneous writing and reading aloud). In the same way, massive shift seems to account for intersensory facilitation when simultaneous stimulation from two or more sensory modalities is of sufficiently low intensity to minimize the extent of the blood-shift gradient used above to explain intersensory rivalry. Reports that the learning of nonsense syllables is enhanced when a hand dynamometer is grasped also may be interpreted in terms of massive shift. Relatively massive forms of blood displacement in all parts of the body are well established for fainting, blushing, postural hypotension, functions of erectile tissue, and for plethysmographic applications to lie detection.

Long palpably clear in peripheral tissues, the role of blood shift within the brain itself is becoming increasingly evident.

REFERENCES

1. ADRIAN, E. D. The electrical activity of the cortex. *Proc. Roy. Soc. Med.*, 1935-6, **29**, 198.
2. CHYATTE, C. Perceptual set and critical flicker frequency. *Amer. J. Opthal.*, 1959, **48**, 352-355.
3. ———. Brain blood-shift theory: A preliminary test through correlations of age with alpha EEG and CFF. *J. of Psychol.*, 1965, **61**, 27-32.
4. HIMWICH, H. E. Brain Metabolism and Cerebral Disorders. Baltimore, Md.: Williams and Wilkins, 1951.
5. INGVAR, D. H. Cortical state of excitability and cortical circulation. In H. H. Jasper, L. D. Proctor, R. S. Knighton, W. C. Noshay, & R. T. Costello (Eds.), *Reticular Formation of the Brain: Symposium.* Boston: Little, Brown, 1958. Pp. 381-412.
6. KÖHLER, W. Zur Theorie des Sukzessvergleichs und der Zeitfehler. *Psychol. Forsch.*, 1923, **4**, 115-175.
7. LORENZO, A. V., FERNANDEZ, C., & ROTH, L. J. Physiologically induced alteration of sulfate penetration into brain. *Arch. Neurol.*, 1965, **12**, 128-132.
8. MACRO, J. A. Auditory-tactual rivalry: A gradient predicted through brain blood-shift theory. Master's thesis, DePaul University, Chicago, Illinois, 1966.
9. NOVAK, S. The effect of electrocutaneous digital stimulation on the detection of single and double flashes of light. *Psychol. Monog.*, 1965, **79**(No. 608), 1-19.
10. RAMON Y CAJAL, S. Algunas conjecturas sobre el mecanismo anatomico de la ideacion. *Revista Medica y Cirugia Practicas*, 1895, **36**, 479-508.
11. RASMUSSEN, T., & PENFIELD, W. Further studies of the sensory and motor cerebral cortex in man. *Fed. Proc. Amer. Soc. Exper. Biol.*, 1947, **6**, 452-560.
12. SCHMIDT, C. F. The Cerebral Circulation in Health and Disease. Springfield, Ill.: Thomas, 1950.
13. SHANNON, R. M. The effects of tactually induced unilateral and bilateral electric shock on the auditory threshold, an interpretation based on brain blood-shift theory. Doctoral dissertation, Loyola University, Chicago, Illinois, 1966.
14. STANLEY, G. The boundary of entopic space. *Psychon. Sci.*, 1966, **5**(2), 45-46.

Editorial Offices
Encyclopaedia Britannica
425 North Michigan Avenue
Chicago 11, Illinois

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 241-247.

THE STUDY OF TARGET SIZE IN PRODUCTION OF BRIGHTNESS, HUE, AND SATURATION FOR INTERMITTENT PHOTIC STIMULATION IN THE RED END OF THE VISIBLE SPECTRUM*[1]

Department of Psychology, Michigan State University

S. HOWARD BARTLEY AND RICHARD J. BALL

A. INTRODUCTION

The present investigation is one of the several involved in the program described quite fully in a recent publication (1). The program has had to do with ascertaining the various sensory end results from using intermittent photic stimulation at subfusional pulse rates. The initial result given extensive study in this program was brightness enhancement. The more recent group of phenomena has pertained to changes in hue and saturation with intermittency as a crucial factor, as various portions of the spectrum were used as stimuli.

One of the unique portions of the spectrum in color theory is the extreme long wavelength band extending beyond, let us say, 650 millimicrons ($m\mu$), for this band is supposed to excite only one type of receptor and, therefore, should result in seeing only a single red hue. In some of the authors' earlier work, it became apparent that this expectation was not borne out. Instead, this waveband resulted in other hues as well.

The present investigation supplies more information on this question as well as reporting the findings from manipulating target area ranging from areas within the fovea to those definitely beyond.

B. APPARATUS AND PROCEDURE

The apparatus utilized has been extensively described previously (2). It consists of a prism monochromator with an episcotister; the assembly involving a variable-sized bipartite field of high spectral purity. One-half of the field is steady and its photic intensity can be varied by a rotary crossed polaroid.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 12, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by Public Health Service Research Grant NB 05260-03, from the National Institute of Neurological Diseases and Blindness.

The other half of the field is made intermittent with the episcotister. The variables manipulated in this particular study are target size, wavelength of target, and pulse-to-cycle fraction (PCF). Intermittency rate is maintained at 9.8 cycles per second.

The methods and procedures of data collection have previously been discussed in detail (2) so that only a basic summation will be given here. The observer views the bipartite field with one-half steady and one-half intermittent but of identical luminance and spectral characteristics. By adjusting the rotary polaroid over the steady half, he matches it in brightness to the intermittent half. Six matches are made alternating ascending and descending approaches. Next, the intermittent target is rated in desaturation in relation to the steady target using a nine-point subjective scale between zero desaturation and total desaturation. Finally, the hue of the intermittent target in relation to the hue of the steady target is judged both in direction and magnitude of shift. Therefore, data are collected on the brightness, saturation, and hue of the intermittent target half, in relation to the steady target half of the bipartite field.

Three observers with normal color vision were utilized. They were presented with six different PCFs ranging from 1/32 to 3/4 and three target sizes, ranging from 6 degrees to 1-1/2 degrees. On one observer[2] a fourth target size of 3/4 degree was also used. These target sizes were chosen so that the 1-1/2 degree and 3/4 degree targets would be strictly foveal, while the 3 degree and 6 degree would definitely be parafoveal. The wavelength bands utilized in this particular study ranged from 650 mμ to 710 mμ. The luminosity was the maximum obtainable for each waveband but was not uniform for all wavebands. The maximal obtainable luminosity progressively decreased from 650 mμ to 710 mμ. At 650 mμ, luminosity was 1200 candles per square foot; at 660 mμ, 830; at 670 mμ, 598; at 680 mμ, 432; and by 710 mμ, it had very markedly decreased.

C. Results

Figure 1 shows the effects on brightness of varying target size, PCF, and wavelength for the two observers. The third observer showed essentially the same results. The abscissa is wavelength and the ordinate scale is brightness index. Brightness index is a scale arrived at by relating the obtained brightness to the expected brightness. For comparing points on any given PCF curve, this represents a measure of absolute differences in perceived

---

[2] R. J. Ball.

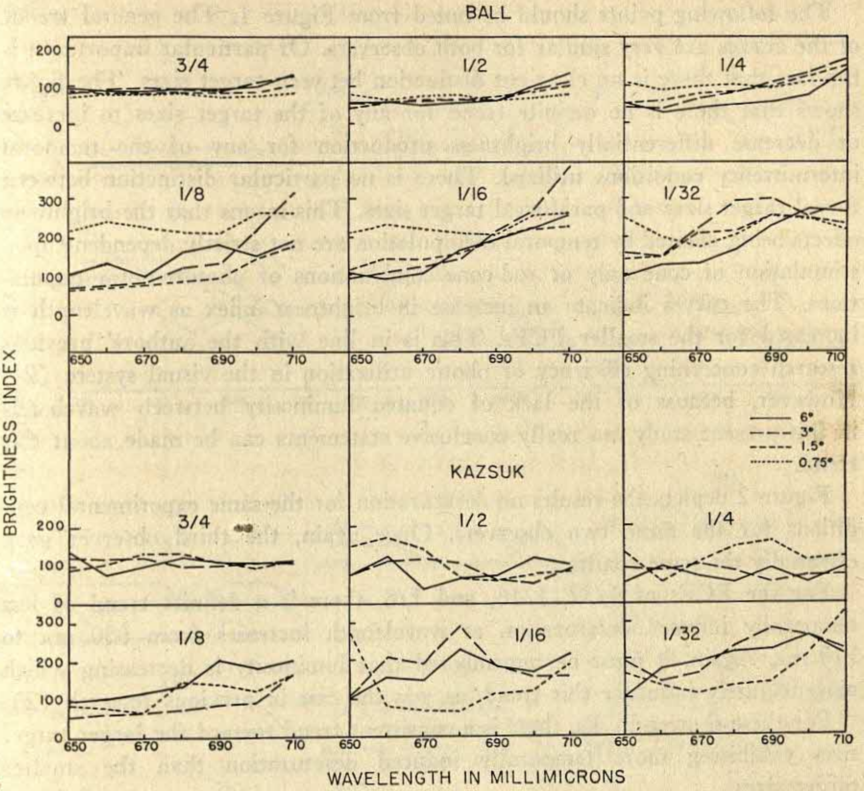

FIGURE 1

THE RELATION BETWEEN WAVELENGTH AND BRIGHTNESS INDEX FOR TARGETS OF VARIOUS SIZES

Each family of curves is for a single PCF. PCFs range from 3/4 to 1/32. Target sizes are indicated in the figure.

brightness, while in relating different PCF curves this represents differences in the ratio of obtained to expected brightness. The rationale for utilizing a brightness index scale has been thoroughly discussed previously (2, 4).

The level 100 is the point where obtained brightness was equal to the expected brightness; 200 is the level at which twice the expected luminance was required in the steady part of the field to match the brightness of the flickering part of the field. The expected brightness is based upon the portion of the intermittency cycle during which the stimulus is presented and is experimentally determined by multiplying the luminance level of the pulse by the pulse-to-cycle fraction (PCF).

The following points should be noted from Figure 1. The general trends of the curves are very similar for both observers. Of particular importance is the fact that there is no clear-cut distinction between target sizes. The figure shows that there is no definite trend for any of the target sizes to increase or decrease differentially brightness production for any of the temporal intermittency conditions utilized. There is no particular distinction between foveal target sizes and parafoveal target sizes. This means that the brightness effects being elicited by temporal manipulation are not strictly dependent upon stimulation of cone only or rod-cone combinations of photoreceptor populations. The curves indicate an increase in brightness index as wavelength is increased for the smaller PCFs. This is in line with the authors' previous research concerning efficiency of photic utilization in the visual system (2). However, because of the lack of equated luminosity between wavebands in this present study, no really conclusive statements can be made about this trend.

Figure 2 depicts the results on desaturation for the same experimental conditions for the same two observers. Once again, the third observer gave essentially the same results.

For the PCFs of 1/32, 1/16, and 1/8, there is a definite trend of less temporally induced desaturation as wavelength increases from 650 mμ to 710 mμ. Again, it must be remembered that luminosity is decreasing which may definitely influence this trend, as was the case in previous research (2).

For these shorter PCFs, there is a consistent trend toward the larger target sizes exhibiting more temporally induced desaturation than the smaller target sizes.

With a PCF of 3/4, there is very little desaturation for any wavelength and no consistent trend relating to target size. The PCF of 1/2 gives somewhat more desaturation than the PCF of 3/4, but no consistent difference based on target size. The PCF of 1/4 gives still more desaturation with the beginning of the trend of larger target sizes giving greater desaturation.

The third kind of data collected was a set concerning temporal induction of hue shift. The agreement among the three observers was again good. Several interesting points should be made concerning these data. Very strong temporal hue shifts were induced under some of the experimental conditions. The hues for the long wavelengths were, in many cases, shifted by using intermittent stimulation to hues normally associated with wavelengths of 550 mμ to 620 mμ. In other words, those wavelengths above 650 mμ were perceived under various temporal manipulations as yellow, yellow-orange, orange, red-orange and, in some instances, even as chartreuse. These are very striking

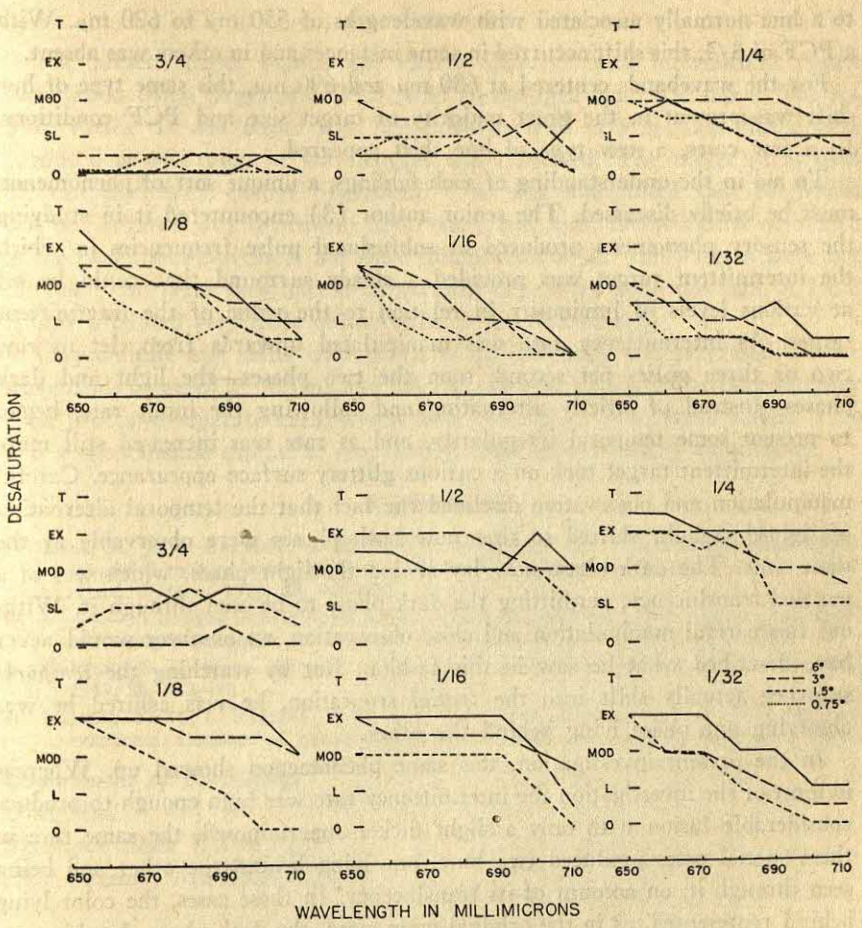

FIGURE 2

The Relation Between Desaturation and Wavelength for Targets of Various Sizes

Each family of curves is for a single PCF. PCFs range from 3/4 to 1/32. Target sizes are indicated in the figure.

and consistent results that strongly diverge from many photochemically based statements that hue should be constant above 650 mμ, and only brightness should change because only one cone pigment is being activated. The authors' findings very definitely do not fit this supposition.

For the wavebands centered at 650 mμ, 660 mμ, and 670 mμ, there was for all target sizes and all PCFs, except 3/4, a temporally induced hue shift to a hue normally associated with wavelengths of 550 mμ to 620 mμ. With a PCF of 3/4, this shift occurred in some instances and in others was absent.

For the wavebands centered at 680 mμ and 690 mμ, this same type of hue shift was present in the great majority of target size and PCF conditions. In a few cases, a new type of hue shift appeared.

To aid in the understanding of such findings, a unique sort of phenomenon must be briefly discussed. The senior author (3) encountered it in studying the sensory phenomena produced at subfusional pulse frequencies in which the intermittent target was provided a steady surround that could be set at various levels of luminosity in relation to the value of the intermittent target. As intermittency rate was manipulated upwards from, let us say, two or three pulses per second, soon the two phases—the light and dark phases—instead of strictly alternating and following the input rate, began to present some temporal irregularity, and as rate was increased still more the intermittent target took on a curious glittery surface appearance. Careful manipulation and observation disclosed the fact that the temporal alternation of light-dark-light shifted so that now both phases were observable at the same time. The dark phase now lay *behind* the light phase, which was of a mottled translucency, permitting the dark phase to be seen through it. Without this careful manipulation and close observation, an observer would never have described what he saw in this fashion. But by watching the *temporal* sequence actually shift into the *spatial* separation, he was assured he was observing one phase lying behind the other.

In the present investigation, this same phenomenon showed up. Whereas in most of the investigation the intermittency rate was high enough to produce considerable fusion with only a slight flicker superimposed, the same rate in these special cases produced two hues, one lying behind the other and being seen through it, on account of its translucency. In these cases, the color lying behind represented, as in the original cases cited, the dark phase. In this case, it was a dark purplish hue. The front or light phase was orange. For the wavebands of 700 mμ and 710 mμ, more conditions gave the purple-orange hue shift, many still gave simply a hue shift to orange, and about half of the conditions gave no hue shift. It must be remembered here, however, that luminosity was at a low level for these wavebands. We are presently constructing equipment to permit investigation of wavebands up to 750 mμ under much higher luminosity conditions.

D. SUMMARY

The present paper reports the effects obtained with intermittent photic stimulation when target area, PCF, and wavelength from 650 mμ to 710 mμ were varied. It was found that hue shifts as well as brightness shifts were produced as this range of wavelengths was traversed. The target areas used did not essentially change brightness. This was interpreted that brightness was not differentially dependent upon cones and rods. Desaturation decreased as the range was traversed, but since luminosities at all wavelengths were not the same, this trend may have been partially dependent upon this fact. A PCF of 3/4 produces very little desaturation at any wavelength.

Very definite hue shifts were produced as the range was traversed. In the wavebands centered at 680 mμ and 690 mμ, the two phases of the intermittency cycle produced separate hues, which were seen to lie in separate planes. This was a phenomenon similar to one earlier found in achromatic stimulation.

REFERENCES

1. BALL, R. J., & BARTLEY, S. H. Further investigations of color perception under temporal manipulation of photic stimulation. *Amer. J. Optom. & Arch. Amer. Acad. Optom.*, 1966, **43**, 419-430.
2. ———. Changes in brightness index, saturation, and hue produced by luminance-wavelength-temporal interactions. *J. Opt. Soc. Amer.*, 1966, **56**, 695-699.
3. BARTLEY, S. H. Some effects of intermittent photic stimulation. *J. Exper. Psychol.*, 1939, **25**, 462-480.
4. NELSON, T. M., & BARTLEY, S. H. The Talbot-Plateau Law and the brightness of restricted numbers of photic repetitions at CFF. *Vision Res.*, 1964, **4**, 403-411.

Department of Psychology
Olds Hall
Michigan State University
East Lansing, Michigan 48823

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 249-255.

SEX DIFFERENCES IN SMOKING ATTITUDE, BEHAVIOR, AND BELIEFS OF COLLEGE STUDENTS*

Department of Psychology, Boston College

DANIEL J. BAER

A. INTRODUCTION

Although more males smoke than females, the reasons for this greater incidence are not fully understood. Recent studies of smokers and nonsmokers (1, 4, 7, 8, 9, 10, 12, 14) have reported personality as well as family background differences of these groups. The changing nature of sex roles in the American culture (3) could contribute to the recent proportionate increase of female smokers (2, 11). However, other studies indicate that constitutional factors (5, 13) are important determiners of the smoking habit.

Because of the close correspondence of attitudes and behavior, it would seem that sex differences in smoking attitude may also account for the greater number of male smokers. The purpose of the present study was to compare college male and female Heavy Smokers, Light-Moderate Smokers, Former Smokers, and Nonsmokers in their smoking attitude, behavior, and beliefs.

B. METHOD

1. *Subjects*

The *S*s were white, American-born Catholic college students ranging from 17 to 23 years in age. Heavy Smokers (males $N = 143$, females $N = 86$) smoked at least a pack of cigarettes a day, or for the males an equivalent amount of pipes and cigars, for at least one month prior to being tested. Light-Moderate Smokers (males $N = 106$, females $N = 73$) actively smoked at the time they were tested, but not as much as Heavy Smokers. Former Smokers (males $N = 32$, females $N = 47$) were at one time habitual smokers but had not smoked for at least one month prior to being tested. Nonsmokers (males $N = 124$, females $N = 73$) had never smoked except for a possible rare trial experience.

2. *Test Instruments*

A Thurstone-type Smoking Attitude Scale, the development of which is reported elsewhere (1), consisted of two sets of 22 statements related to

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 15, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

smoking attitude. The sum of the two median weights of those items to which the *S* checked agreement represented his score. Low values indicated a favorable and high values an unfavorable smoking attitude.

For the Heavy Smokers and Light-Moderate Smokers, a Smoking Behavior and Belief Questionnaire was developed that consisted of 32 questions dealing with their smoking activities as well as other biographical information. Each question required a "yes" or "no" response. A similar set of 30 questions was presented to the Former Smoker group. Because of their limited experiences with tobacco, only 11 questions were offered to the Nonsmokers. Nine questions were the same on the three Smoking Behavior and Belief Questionnaires.

3. *Procedure*

The *S*s were secured from many sources on campus by 20 male and 15 female students. Male examiners tested only male, while female examiners tested only female *S*s. Examiners first administered the Smoking Attitude Scale, then classified the *S* according to his smoking behavior, and finally had the *S* complete the appropriate Smoking Behavior and Belief Questionnaire.

C. RESULTS AND DISCUSSION

Table 1 summarizes the mean and standard deviation of the smoking attitude of the male and female Heavy Smokers, Light-Moderate Smokers, Former Smokers, and Nonsmokers. The distribution of scores for the groups was approximately normal in shape, while the standard deviations were comparable.

Results of the *t*-test of the difference of the mean smoking attitude of each group compared to each of the other groups are reported in Table 2. It may be seen that males differ from females for each of the four groups, with the males consistently having the more favorable smoking attitude.

TABLE 1
MEAN AND STANDARD DEVIATION OF SMOKING ATTITUDE OF MALE AND FEMALE HEAVY SMOKERS, LIGHT-MODERATE SMOKERS, FORMER SMOKERS, AND NONSMOKERS

Group	Sex	*N*	Mean	*SD*
Heavy smokers	Male	143	8.6	3.2
	Female	73	9.3	2.4
Light-moderate smokers	Male	106	8.7	2.4
	Female	86	9.6	2.3
Former smokers	Male	32	10.4	3.0
	Female	47	11.8	2.7
Nonsmokers	Male	124	11.9	2.5
	Female	73	13.2	3.3

TABLE 2

t-TEST OF THE DIFFERENCE BETWEEN MEAN SMOKING ATTITUDE OF MALE AND FEMALE HEAVY SMOKERS, LIGHT-MODERATE SMOKERS, FORMER SMOKERS AND NONSMOKERS

Group	Female heavy smokers	Male light-moderate smokers	Female light-moderate smokers	Male former smokers	Female former smokers	Male nonsmokers	Female nonsmokers
Male heavy smokers	2.05*	0.47	2.83**	3.87**	7.48**	9.78**	12.71**
Female heavy smokers		1.51	0.72	2.08*	4.91**	6.03**	8.85**
Male light-moderate smokers			2.20*	3.34**	6.60**	8.39**	11.26**
Female light-moderate smokers				1.28	3.86**	5.16**	7.57**
Male former smokers					2.03*	2.42**	2.67**
Female former smokers						0.17	2.64**
Male nonsmokers							3.07**

* $p < .05$ (two-tailed hypothesis).
** $p < .01$ (two-tailed hypothesis).

Thus, not only did male Heavy Smokers report a significantly (.01 level) more favorable smoking attitude than did female Heavy Smokers, but similar differences occurred between male and female Former Smokers and Nonsmokers. This more favorable attitude of the males may not only be due to their less inhibited or more exaggerated responses, but it also may reflect real differences in attitude about smoking. Parental training, peer group influences, social expectancy, as well as prior smoking experiences probably discourage females from smoking more than they discourage males, and promote in females a more unfavorable smoking attitude.

It may be seen from Tables 1 and 2 that, regardless of sex, as smoking experience increases, there is a correspondingly more favorable smoking attitude. Also, the smoking attitude of both male and female Former Smokers is less favorable than that of Heavy Smokers and Light-Moderate Smokers, and yet is more favorable than that of Nonsmokers. The question of whether or not attitudes produce or are the result of smoking behavior cannot be answered from the present study. However, since at one time everyone was a nonsmoker, it would be interesting to study longitudinally the smoking attitude of individuals. In this way, a better understanding of the reasons for smoking, not smoking, or discontinuing smoking could be established.

Although male Heavy Smokers and Light-Moderate Smokers were not different, these two groups held a significantly more favorable smoking attitude than did male Former Smokers and Nonsmokers, while male Former Smokers reported a more favorable attitude than did Nonsmokers. This same outcome occurred when the four female groups were compared to one another. Such a consistency of smoking attitude and smoking behavior would be expected if attitudes truly reflect behavior. Since the differences in smoking behavior between Heavy Smokers and Light-Moderate Smokers are quantitative rather than qualitative, a smoking attitude difference would seem to be less likely for these groups than when these groups were compared to others who do not smoke.

When the dissimilar groups were compared, male Heavy Smokers indicated a significantly (.01 level) more favorable attitude than did female Light-Moderate Smokers, Former Smokers, and Nonsmokers, while female Heavy Smokers reported a significantly less favorable smoking attitude than did male Light-Moderate Smokers, but significantly more favorable attitude than did Nonsmokers. It would seem that female Heavy Smokers maintain this habit for reasons different from those of male Heavy Smokers. Perhaps this outcome may be partly due to the definition of a heavy smoker as one who smokes 20 or more cigarettes a day. Females may light that number of cigarettes, but the total number of puffs on these cigarettes may be less for the females than for the males.

For each of the nine questions common to the three Smoking Behavior and Belief Questionnaires, 2 × 2 contingency tables were formed, comparing the number of males and females who responded "yes" and "no." In Table 3 are reported the percentage of "yes" responses for each sex and the chi square test of independence of the responses for the four most discriminating questions. Because the percentage of the "yes" responses were similar for Heavy Smokers and Light-Moderate Smokers, and because the expected values for some cell frequencies were less than five, the two smoking groups were combined into the one group identified as Present Smokers.

TABLE 3
SEX DIFFERENCES IN SMOKING BELIEFS OF PRESENT SMOKERS, FORMER SMOKERS AND NONSMOKERS

Question	Group	Male % yes	Female % yes	Chi square
Do you believe there is a relationship between smoking and lung cancer?	Present smokers	67.1	88.1	22.93**
	Former smokers	100.0	93.6	2.85
	Nonsmokers	74.2	97.3	17.12**
Do you believe there is a relationship between smoking and reduction in life expectancy?	Present smokers	45.0	62.9	12.47**
	Former smokers	62.5	81.0	6.00**
	Nonsmokers	58.1	76.7	7.02**
Does the smoke from smoking bother you?	Present smokers	16.5	28.5	5.25*
	Former smokers	46.9	46.8	0.01
	Nonsmokers	53.2	68.5	4.42*
Would you want your own children to smoke?	Present smokers	17.3	4.4	14.90**
	Former smokers	6.2	2.1	0.80
	Nonsmokers	8.9	0.0	6.90**

* $p < .05$.
** $p < .01$.

For the question, "Do you believe there is a relationship between smoking and lung cancer?" 67.1 per cent of the males and 88.1 per cent of the females of the Present Smokers responded "yes." The significant chi square (.01 level) indicated that the proportion of males differed from the proportion of females in their "yes" responses. For the same question, significant sex differences were observed for the Nonsmokers, but not for the Former Smokers.

In general, the results of Table 3 indicate that regardless of sex, Present Smokers were less inclined to believe in the dangers to health from smoking, were less bothered by the smoke from smoking, and were more willing to allow their future children to smoke, than were Former Smokers and Nonsmokers.

However, for each group proportionately, more females than males were concerned with the harmful effects of smoking. Since the direction of these differences corresponds to the attitude differences of Table 2, it would seem that the smoking attitude is consistent with the smoking behavior and practices of both the male and female groups.

Since proportionately more male and female Former Smokers than Present Smokers indicate that they believe there is a relationship between smoking and lung cancer and reduction in life expectancy, results of the present study do not support the conclusion of Weatherley (14), that concern with health is not necessarily associated with the actual stopping of smoking. The lung-cancer scare seems to have uniformly impressed the male Former Smokers, while a more general concern with the reduction in life expectancy and smoking has influenced both male and female Former Smokers.

When male and female Present Smokers were compared in their responses to the Smoking Behavior and Beliefs Questionnaire, several significant differences were observed. The chi squares in Table 4 indicate that a greater percentage of males were Heavy Smokers, liked the taste of smoking, received pleasure from smoking, and began smoking before entering high school or college, than females. However, more females would like to stop smoking, believed they smoked too much, and preferred filter-tip and king-size cigarettes than did males. These results indicate that female smokers not only differ from male smokers in the type of smoking experiences, but also in satisfaction received from smoking. If the motivation for females to continue smoking is different from that of the males, then the procedures necessary to encourage and promote the discontinuance of smoking for each sex should allow for these differences.

TABLE 4
SEX DIFFERENCES IN SMOKING BEHAVIOR AND BELIEFS OF PRESENT SMOKERS

Question	Males % yes ($N = 249$)	Females % yes ($N = 159$)	Chi square
Are you a heavy smoker?	57.4	45.9	5.17*
Would you like to stop smoking?	34.1	59.1	24.60**
Do you believe you smoke too much?	48.2	60.4	5.78*
Do you like the taste from smoking?	77.5	58.5	16.75**
Do you receive any pleasure from smoking?	96.8	89.9	8.22**
Do you try to limit the amount you smoke?	47.0	70.4	21.70**
Do you prefer filter cigarettes?	37.3	97.0	143.80**
Do you prefer king-size cigarettes?	42.6	70.4	30.30**
Did you smoke before entering high school?	17.3	6.9	9.05**
Did you smoke before entering college?	80.3	64.8	12.26**

* $p < .05$.
** $p < .01$.

D. SUMMARY

In order to determine the sex differences in smoking attitude, behavior, and beliefs, a sample of 405 male and 279 female college students were administered a Thurstone-type Smoking Attitude Scale and a Smoking Behavior and Belief Questionnaire. When Heavy Smokers, Light-Moderate Smokers, Former Smokers, and Nonsmokers were compared in their smoking attitude, it was found that, regardless of sex, the greater the smoking experience, the more favorable the smoking attitude. However, when matched for smoking behavior, males reported a more favorable smoking attitude than did females. Although Present Smokers were less inclined to believe in the dangers to health from smoking than were Former Smokers and Nonsmokers, more females than males were concerned with these effects when groups of similar smoking experience are compared.

REFERENCES

1. BAER, D. J. Smoking behavior and beliefs of college males. *J. Soc. Psychol.*, 1966, **68**, 65-78.
2. BONNELL-LEWIS, W. H. B. The smoking habits of women students in a teachers' training college: Pilot survey. *Health Educ. J.*, 1963, **21**, 4-11.
3. BROWN, D. G. Sex role development in a changing culture. *Psychol. Bull.*, 1958, **55**, 232-242.
4. CARTWRIGHT, A., & THOMSON, J. G. Young smokers: An attitude study among school children touching upon parental influence. *Brit. J. Prev. & Soc. Med.*, 1960, **14**, 28-34.
5. EYSENCK, H. J. Personality and cigarette smoking. *Life Sci.*, 1964, **3**, 777-792.
6. ———. Smoking, Health, and Personality. New York: Basic Books, 1965.
7. LAWTON, M. P. Psychosocial aspects of cigarette smoking. *J. Health & Hum. Behav.*, 1962, **3**, 163-170.
8. PHANISHIGI, R. A. Causes of smoking habit in college students—An investigation. *J. Educ. & Psychol., Barvda*, 1951, **9**, 29-37.
9. SALBER, E. J., & ROCHMAN, J. E. Personality differences between smokers and nonsmokers. *Arch. Environ. Health*, 1964, **8**, 459-465.
10. SALBER, E., & WORCESTER, J. Changes in women's smoking patterns. *Cancer*, 1964, **17**, 32-36.
11. SALBER, E. J., WELSH, B., & TAYLOR, S. V. Reasons for smoking given by secondary school children. *J. Health & Hum. Behav.*, 1963, **4**, 118-129.
12. SCHUBERT, D. S. P. Impulsivity and other personality characteristics of cigarette smokers. *Clin. J. Soc. Psychiat.*, 1964, **10**, 305-308.
13. TODD, G. F. Concordance of smoking habits in monozygotic and dizygotic twins. *Heredity*, 1959, **13**, 417-444.
14. WEATHERLEY, D. Some personality correlates of the ability to stop smoking cigarettes. *J. Consult. Psychol.*, 1965, **29**, 483-485.

Department of Psychology
Boston College
Chestnut Hill, Massachusetts 02167

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 257-264.

THE EFFECTS OF PHENOTHIAZINE DRUGS ON VERBAL CONDITIONING OF SCHIZOPHRENICS*

University of Minnesota Medical School[1]

NORRIS D. VESTRE

A. INTRODUCTION AND PURPOSE

One of the more objective of the many verbal conditioning techniques, originated by Taffel (10), sets up a conditioning situation in which subjects are asked to make up sentences with a given verb and a choice of one of six pronouns. Whenever subjects begin a sentence with one of two pronouns selected as critical responses (usually I and We) the experimenter says "good" or "mm-hm" in an attempt to reinforce these particular pronouns.

Although conditioning by this procedure has been demonstrated in various types of populations, the responsivity of schizophrenics to conditioning is unclear. Cohen and Cohen (2), who found that schizophrenics were not conditionable, indicate that this would seem to support the hypothesis that schizophrenics are deficient in responsiveness to social cues. Leventhal (6) found that schizophrenics did learn under punishment-only or combined reward and punishment conditions, but not under reward-only conditions (i.e., to "good"). Ebner (3) also found that schizophrenics failed to show the conditioning effect. Conversely, Hartman (5) and Ells (4) report that certain of their schizophrenic groups were conditioned by "good."

Although schizophrenia may involve a deficient responsivity to social stimuli, another possible explanation of these negative findings in verbal conditioning studies of schizophrenics is the effect of "tranquilizing" drugs. Most investigators have failed to indicate the drug status of their subjects. In view of the wide use of phenothiazines in the treatment of hospitalized psychiatric patients, it seems likely that the schizophrenics used in most verbal conditioning studies have been receiving phenothiazines. There is a greater probability that the schizophrenics used by Hartman (5) were drug-free or mostly

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 16, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The data on which this article is based were collected while the author was at the Veterans Administration Hospital, St. Cloud, Minnesota. The study was supported, in part, by the Veterans Administration Medical Research Program Funds and by Research Grant MYP-5106, National Institute of Mental Health, United States Public Health Service.

drug-free, since his research was done about the time that the phenothiazines were introduced, before they became as widely used as they are now. Ells (4), the other investigator to condition schizophrenics using "good," limited his schizophrenics to those on low doses of drugs: *viz.,* a maximum of 200 mg chlorpromazine or its equivalent in other drugs. Only one investigation of the effects of psychoactive drugs on conditionability in the verbal conditioning situation has been reported. In that study (11) it was found that schizophrenics being treated with one of two phenothiazines were not conditioned, whereas schizophrenics receiving phenobarbital were conditioned. The absence of a placebo control group in that study made it impossible to infer what a drug-free schizophrenic would do in the conditioning situation.

The wide use of phenothiazines with schizophrenic patients makes it important to explore further the effects of these drugs on susceptibility to verbal conditioning. This investigation compares the effects of three phenothiazines —thioridazine, trifluoperazine, and chlorpromazine—and of placebo on susceptibility to verbal conditioning among schizophrenics in the Taffel-type conditioning procedure.

B. METHOD

1. *Subjects and Drug Procedures*

All subjects were males and ranged from 29 to 55 years in age, the mean being 41. All had a diagnosis of schizophrenia with no history or other evidence of complicating organic factors. The length of the present hospitalization for the group ranged from one year to over 18 years, the mean duration being 10 years.

The subjects in the present study were participating in another study (8)[2] that was designed to evaluate the effectiveness of three different drugs in reducing psychiatric symptomatology. In this study, 80 hospitalized schizophrenics were randomly assigned to one of four treatments (thioridazine, trifluoperazine, chlorpromazine, or placebo) given in identical-appearing capsules. The thioridazine and chlorpromazine capsules contained 100 mg each; the trifluoperazine capsules contained five mg. The psychiatrist was permitted to adjust dosages between two and 10 capsules per day in order to obtain maximum benefit for each patient. In the case of trifluoperazine, this permitted a range of 10 to 50 mg per day, and for thioridazine and chlorpromazine, a range of 200 to 1000 mg per day.

The subjects available for the verbal conditioning study consisted of those patients still on medication after six weeks. During that time, one patient had been lost from the placebo group but none had been lost from the three drug groups. Patients within each of these four treatment groups were assigned on a random basis to an experimental or a control group. At the time of the conditioning trials the experimenter did not know which drug any patient was receiving.

---

[2] The drugs used in this study were supplied by Sandoz Pharmaceuticals and by Smith Kline and French Laboratories.

2. *Verbal Conditioning Procedure*

The materials consisted of 80 three-by-five-inch index cards, on each of which was typed the simple past tense of a commonly used verb and six pronouns (I, WE, YOU, HE, SHE, THEY). Eighty of the 720 possible different orders of pronouns were randomly assigned to the 80 cards.

Subjects were shown a sample card and given instructions to make up a sentence beginning with any one of the pronouns and containing the verb. In order to determine operant levels (i.e., baseline frequencies) for the various pronouns, none of the subjects was given reinforcement during the first 20 trials (cards). Control subjects were treated the same during Trials 21-80 as they were for Trials 1-20: i.e., the experimenter said nothing as they were making up sentences. For the experimental subjects, Trials 21-80 were conditioning trials; the experimenter verbalized a soft, unemotional "good" following sentences beginning with criterion pronouns, and said nothing following sentences beginning with noncriterion pronouns. The selection of the pronouns to reinforce was based on the operant levels for the various pronouns. In the case of each subject, those two pronouns which ranked third and fourth in frequency (the two middle-most ranks) during operant trials were selected as criterion responses.

At the completion of 80 trials, each subject was asked the following questions:

1. What did you think this was all about?
2. How did you decide which of the words on the bottom to use?
3. Which do you think you used most often?
4. Why?
5. Was I doing anything while you were making up sentences?
6. Did you hear me say anything?
7. Did you hear me say "good" while you were making up sentences?
8. What did you think it meant?
9. When did I say it?

If the subject reported, in response to one of the first six questions, that he heard the experimenter say "good," he was asked Questions 8 and 9 to determine if he could verbalize a contingency between his responses and the experimenter's utterances. Subjects who said "No" to Question 6 were asked Question 7. Subjects who said "No" to Question 7 were questioned no further; those who said "Yes" were asked Questions 8 and 9.

C. RESULTS

Of the 79 patients still on medications after six weeks and available for the conditioning trials, 30 had to be dropped. Twenty-two of these 30 could not do the task because they either failed to understand the directions or read poorly and required help pronouncing the verbs. Since criterion pronouns were determined on the basis of operant level performance, it was required that at least four different pronouns be used during Trials 1-20, and four subjects who failed to do this were dropped. One subject was lost because of the experimenter's error (i.e., reinforcing a control subject) and three subjects refused to participate. None of the subjects verbalized the contingency between their responses and the experimenter's verbalizations. These dropouts reduced the trifluoperazine group to five experimental and six control subjects, the chlorpromazine group to 10 experimental and five control subjects, the thioridazine group to seven experimental and five control subjects and the placebo group to seven experimental and four control subjects.

None of the differences among the means for age or length of hospitalization was found to be statistically significant. Any between-group differences in conditionability, therefore, cannot be accounted for simply on the basis of these variables.

For statistical purposes the data were divided into four blocks of 20 trials each, with Block 1 being the operant level. Each of the three blocks of conditioning trials was analyzed by an analysis of covariance with the Block 1 (operant level) scores being used as the covariate to adjust the scores on Blocks 2, 3, and 4. Since there were unequal and disproportionate frequencies in the subgroups, the method of unweighted means (9) was used.

The analysis of covariance for Block 2 (Trials 21-40) yielded no evidence of any conditioning effect, drug effect, or interaction (treatment × drug) effect.

For Block 3 (Trials 41-60) a significant (.05 level) *F* ratio was obtained between treatments: i.e., experimental *vs.* control. For Block 4 (Trials 61-80) a significant (.05 level) *F* ratio was obtained between treatments. These results are shown in Tables 1 and 2. For Block 4, the *F* ratio testing the interaction (treatment × drug) effect was significant at the .10 level and almost reached the .05 level. An *F* ratio of 2.74 was obtained and 2.84 is the

TABLE 1
ANALYSIS OF COVARIANCE FOR BLOCK 3 TRIALS
(Adjusted on Operant Level)

Source of variation	Degrees of freedom	Adjusted sum of squares	*F*
Treatment	1	7.70	5.83*
Drug	3	1.25	.32
Interaction	3	3.15	.80
Error	40	52.65	

* .05 level.

TABLE 2
ANALYSIS OF COVARIANCE FOR BLOCK 4 TRIALS
(Adjusted on Operant Level)

Source of variation	Degrees of freedom	Adjusted sum of squares	*F*
Treatment	1	9.83	7.28*
Drug	3	7.95	1.96
Interaction	3	11.10	2.74
Error	40	53.85	

* .05 level.

level with $df = 3, 40$. Scheffe's test (7) was used to test the significance of differences between the adjusted means of individual pairs of groups. Each experimental group was compared with its control group and each of the four experimental groups was compared with each other. None of these comparisons was significant. The extremely small sample sizes in these comparisons, of course, tended to make it unlikely that significance would be reached. The placebo control group, for example, had only four subjects.

As a further check on the conditioning performance of the four experimental groups, it was decided to combine the control groups. This would permit separate comparisons between the means of the experimental groups and the combined control group. For ease of analysis the scores of the last two blocks of conditioning trials were summed and the operant levels were subtracted from this total. The results of *t*-tests computed between the mean of each experimental group and the mean of the combined control group are shown in Table 3. Since the assumption of equal variances seemed untenable, the method of adjusting degrees of freedom described in Walker and Lev (12, pp. 157-158) was used. Significant (.05 level) *t* values were obtained for the thioridazine and trifluoperazine groups. Although the *t* value for the placebo group only reached the .10 level, it was very close to the .05 level (a *t* value of 2.23 was required). The *t* value for the chlorpromazine group

TABLE 3

SUMMARY OF t-TESTS FOR EXPERIMENTAL (E) GROUPS *vs.* COMBINED CONTROL GROUP

Group	N	Mean	Variance	t
Thioridazine E	7	10.57	32.95	2.61*
Trifluoperazine E	5	7.80	4.70	2.54*
Chlorpromazine E	10	6.70	37.12	1.06
Placebo E	7	10.57	49.62	2.18**
Combined control	20	4.45	15.94	

Note: t-values are comparisons of the experimental (E) group with the combined control group.

* .05 level.

** .10 level.

was not significant. Although this method of analysis is admittedly somewhat crude, it does give some further evidence of the performance of the four experimental groups in the conditioning situation.

D. DISCUSSION

This investigation was designed to explore the effects of tranquilizing drugs on conditionability of verbal behavior in schizophrenics. The data from the study suggested that the verbal behavior of schizophrenics receiving thioridazine, trifluoperazine, or placebo was modified. The evidence for conditioning in the placebo group, however, was not as strong as it was in the drug groups. The verbal behavior of the chlorpromazine-treated subjects, on the other hand, was not affected by the verbal conditioning procedure.

Design limitations tend to complicate these conclusions to some extent. First, a large number of patients were unable to cooperate in the conditioning trials. Twenty-two subjects were dropped for this reason and an additional three subjects refused to participate. This is perhaps evidence of the extreme morbidity and chronicity of the sample selected. However, approximately the same number of patients were lost from each group by reason of refusal or by reason of inability to do the task: *viz.,* eight placebo subjects, six thioridazine subjects, six trifluoperazine subjects, and five chlorpromazine subjects. The major problem of this investigation was the sizeable sample attrition that led to a reduction of the power of the statistical tests between pairs of groups. This, of course, was the justification for combining the control groups in order to explore further the drug effects on conditioning.

Although the drug dosages were thought to be comparable, they may not have been. A flexible dosage schedule was permitted whereby the psychiatrist could adjust dosages between specified ranges. At the time of the conditioning trials, there were no significant between-group differences in the daily number of capsules administered. The mean daily dosages actually received at the time of conditioning trials were as follows: trifluoperazine, 5.29 capsules (26.45 mg); chlorpromazine, 5.79 capsules (579 mg); thioridazine, 6.00 capsules (600 mg); and placebo, 6.68 capsules.

With the assumption that these results could be replicated in another study with larger samples, it is interesting to speculate on the suggested differences between drugs with respect to response to conditioning. Patients receiving placebo showed a conditioning effect indicating that schizophrenics can be conditioned with positive reinforcers. Although patients receiving chlorpromazine showed no conditioning, patients receiving trifluoperazine or thioridazine did show conditioning. Trifluoperazine, in contrast to chlorpromazine, is reputed to have a stimulating type of action. Antidepressive activity is claimed for both thioridazine and trifluoperazine, but not for chlorpromazine (1). Perhaps this is the basis for the between-drug differences in responsivity to conditioning found in this study.

The findings of the present study would seem to suggest one reason why other investigators [e.g., Cohen and Cohen (2)] have been unable to condition schizophrenics. Certain phenothiazines seem to decrease responsivity to verbal conditioning. Chlorpromazine was the first phenothiazine to be synthesized, is by far the best known, and is still regarded by many psychiatrists as one of the most effective drugs available. For these reasons it has been, by far, the most widely used phenothiazine in a hospitalized population. Consequently, if one were to select a group of hospitalized schizophrenics for a verbal conditioning study, it would be very likely (*a*) that almost all would be receiving phenothiazines, and (*b*) that most of these would be receiving chlorpromazine. At the very least, the data from the present study indicate that drug status is an important variable in accounting for conflicting results in verbal conditioning studies involving hospitalized schizophrenics.

E. SUMMARY

A sample of 79 hospitalized schizophrenics were tested in the verbal conditioning situation after six weeks on one of four treatments (thioridazine, trifluoperazine, chlorpromazine, or placebo). Thirty subjects were not testable for one of a variety of reasons, leaving 49 subjects on whom the data could be used.

Within each treatment group, subjects were assigned to an experimental or a control group. For experimental subjects, the experimenter said "good" whenever the subject started a sentence with a criterion pronoun.

Analysis of criterion pronoun frequencies indicated a significant (.05 level)

conditioning effect for the four groups combined. Comparison of each experimental group with the combined control group revealed that conditioning was demonstrated in the thioridazine, trifluoperazine, and placebo groups, but not in the chlorpromazine group.

REFERENCES

1. BENSON, W. M., & SCHIELE, B. C. Tranquilizers and Antidepressant Drugs. Springfield, Ill.: C C Thomas, 1959.
2. COHEN, E., & COHEN, B. D. Verbal reinforcement in schizophrenia. *J. Abn. & Soc. Psychol.,* 1960, **60**, 443-446.
3. EBNER, E. Verbal conditioning in schizophrenia as a function of degree of social interaction. Unpublished Doctoral dissertation, Purdue University, Lafayette, Indiana, 1961.
4. ELLS, E. M. Verbal conditioning of chronic schizophrenics as a function of operant level and experience with the experimenter. Unpublished Doctoral dissertation, University of Minnesota, Minneapolis, Minnesota, 1963.
5. HARTMAN, C. H. Verbal behavior of schizophrenic and normal subjects as a function of types of social reinforcement. Unpublished Doctoral dissertation, University of Iowa, Iowa City, Iowa, 1955.
6. LEVENTHAL, A. M. The effects of diagnostic category and reinforcer on learning without awareness. *J. Abn. & Soc. Psychol.,* 1959, **59**, 162-166.
7. SCHEFFE, H. The Analysis of Variance. New York: Wiley, 1959.
8. SCHIELE, B. C., VESTRE, N. D., & STEIN, K. E. A comparison of thioridazine, trifluoperazine, chlorpromazine and placebo: A double-blind controlled study on the treatment of chronic schizophrenic patients. *J. Clin. & Exper. Psychopathol.,* 1961, **22**, 151-162.
9. SNEDECOR, G. W. Statistical Methods (5th ed.). Ames, Iowa: Iowa State Coll. Press, 1956.
10. TAFFEL, C. Anxiety and the conditioning of verbal behavior. *J. Abn. & Soc. Psychol.,* 1955, **51**, 496-501.
11. VESTRE, N. D. Relative effects of phenothiazines and phenobarbital on verbal conditioning of schizophrenics. *Psychol. Rep.,* 1965, **17**, 289-290.
12. WALKER, H. M., & LEV, J. Statistical Inference. New York: Holt, 1953.

Department of Psychiatry and Neurology
Medical School
University of Minnesota
Minneapolis, Minnesota 55455

Published as a separate and in *The Journal of Psychology*, 1966, **64**, 265-276.

MALE-FEMALE PERCEPTIONS OF THE FEMALE ROLE IN THE UNITED STATES*

Maferr Foundation and the City College of the City University of New York

ANNE STEINMANN AND DAVID J. FOX

A. INTRODUCTION

For some time now, there has been considerable discussion in both lay and professional circles about the problems faced by women in our changing society. Many authorities of varying degrees of expertness have voiced their opinions on this subject. However, these discussions, although often quite heated and involved, have been remarkably data-free. The current research was undertaken to obtain empirical data on the perception of the female role in terms of two basic questions: first, how do women view themselves, considering themselves as they are, as they would like to be, and as they think men would like them to be; and second, how do men see their ideal woman?

Two hypotheses were tested. The first hypothesis was that despite differences in socioeconomic class, ethnic or racial background, level of education, occupational or professional status, American women share certain values in relation to their roles in and out of the family, and that, specifically, they share a perceived conflict between the level of activity and independence they would like and the much lower level of activity they believe men would prefer them to have. The second hypothesis was that American men verbalize a level of activity they desire for women not different from that level of activity that women say they perceive for themselves.[1]

B. SUBJECTS

This paper is based on data from 10 cluster samples totalling 837 American women, and six cluster samples totalling 423 American men. The samples (indicated in Table 1) include college undergraduates, physicians, lawyers, artists, working women and their husbands, nurses, and Negro professional men and women. Respondents were both Negro and white, and the non-college respondents had all received a high-school education. The age range

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 19, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Current extensions of the study are intended to test these hypotheses in countries other than the United States.

TABLE 1
MEANS AND STANDARD DEVIATIONS FOR SELF-PERCEPTION, IDEAL WOMAN, AND MAN'S IDEAL WOMAN (FOR FEMALE RESPONDENTS); AND MAN'S IDEAL WOMAN (FOR MALE RESPONDENTS), BY SAMPLE

			Female respondents						Male respondents	
	Number		Self-perception		Own ideal		Man's ideal woman		Ideal woman	
Sample	F	M	$\overline{X}$[a]	*sd*	$\overline{X}$	*sd*	$\overline{X}$	*sd*	$\overline{X}$	*sd*
Undergraduates										
Public	96	75	+1.0	11.38	+4.6	13.82	—20.3	14.67	+9.3	8.22
Private (moderate cost)	51	*	+1.5	9.92	+3.6	14.43	—17.4	14.41	*	*
Private (expensive)	68	*	—6.6	11.13	—6.2	14.32	—22.8	14.17	*	*
Physicians	46	50	+5.0	7.81	+3.6	8.23	—13.8	15.64	+2.2	9.49
Lawyers	83	101	+7.6	12.32	+4.5	14.79	—23.0	16.63	—1.0	14.73
Artists	56	53	+8.6	11.03	+5.2	14.61	—16.8	18.15	+3.8	7.91
Business women	53	*	+3.4	10.50	+0.3	12.36	—15.8	17.22	*	*
Members of philanthropic organizations	65	64	+2.1	5.69	**	**	—13.9	12.01	—1.3	11.87
Nurses	218	*	+2.9	5.23	+2.2	10.32	—19.8	15.36	*	*
Negro professionals	91	80	+6.7	12.03	+0.9	16.26	—22.8	17.44	+4.1	13.72
Total	827	423								

[a] Other-achieving (intrafamily) = minus (—); self-achieving (extrafamily) = plus (+).
* No male sample.
** No data available from this sample for this form.

among both female and male samples was wide—ranging from the late teens to the seventies—although a majority were under 40. At the time these data were collected all respondents were residents of New York City, but many of the graduate students were temporary residents and among the professionals several had taken up residence within the past five years. In terms of birthplace, the respondents represented the several geographic regions of the United States, with the largest proportion born in the Northeast, Midwest, and Southeast. While religion specifically was not requested on the Personal Data form used, informal observation indicates that the samples included a reasonable cross section of Catholics, Jews, and Protestants. These were cluster samples selected because of availability and are not presented as random samplings of any groupings within the culture.

C. PROCEDURE

The instrument used for this study is the Inventory of Feminine Values. The Inventory consists of 34 statements, each of which expresses a particular value or value judgment related to women's activities and satisfactions. The respondent indicates the strength of her or his agreement or disagreement to each statement on a five-point scale, ranging from "completely agree" to "completely disagree," through the midpoint of "I have no opinion."

Seventeen of the 34 items are considered to provide a respondent with the opportunity to delineate a family-oriented woman who sees her own satisfactions coming second after those of her husband and family, and who sees her family responsibilities as taking precedence over any potential personal occupational activity. Marriage, child-rearing, and succorant nurturing are the main avenues of this woman's achievement. The other 17 items delineate a self-achieving woman who considers her own satisfactions equally important with those of her husband and family and wishes opportunities to realize any latent ability or talent that she might have. The score on the Inventory represents the difference in strength of agreement to the intrafamily and extrafamily items. A respondent who took identical but opposite positions would have a score of zero; a respondent who consistently took the strongest possible passive position would have a score of —68, and a score of +68 if the strongest possible active position was taken. Positive scores between zero and +68 represent intermediate degrees of an active, self-achieving orientation, while negative scores between zero and —68 represent intermediate degrees of a group or family orientation.

The reliability of the Inventory has been estimated through the split-half technique, and when corrected through the Spearman-Brown procedure is .81.

The items have face validity in that they are generally accepted connotations, but for further verification they also have been submitted to validation by seven judges who were asked to verify the classification of each item as intra- or extrafamily oriented. All items included in the current Inventory were unanimously agreed on by the judges.

Three forms of the Inventory were used in previously reported research with females (1). Each form included the same 34 items in varied order. Female subjects were asked to respond to the items in terms of how they themselves felt (Form SP), how their own ideal woman felt (Form WOI), and in terms of man's ideal woman; i.e., how they thought men would want women to respond (Form WMI).

In the research for this paper on male perception of the female role, male subjects were given a form of the Inventory that asks them to respond to the items as their ideal woman would (Form MI). Overall data from both women and men will be presented first; then the items will be organized into clusters of woman's satisfactions and activities and examined from each of four areas: (*a*) work and accomplishment, (*b*) marriage, (*c*) child-rearing, and (*d*) characteristics of the self.

D. Results

As noted, the data to test the first hypothesis about the homogeneity of female perceptions have already been reported (1). Summarized briefly, these data support the first hypothesis: women *did* share a set of values. Although there was a variety of response within each sample, the average response pattern was the same: most women delineated a self-perception relatively balanced between strivings of self-realization, development, and achievement through their own potentialities—that is, extrafamily strivings—and self-realization and fulfillment by indirection, through permissive nurturing and other-achieving or intrafamily strivings. Women's ideal woman generally was similar to their self-perception, also balanced between self-achieving and other-achieving strivings. Their ideal was slightly more active than the self. However, their perception of man's ideal woman was a woman with little of the self-assertion and self-achievement they reported in themselves or in their ideal woman. Women saw man's ideal woman as significantly more accepting and permissive than their own self-perception of a subordinate role in both personal development and women's place in the familial structure.

Table 1 presents, for review, these data from the samples of women. In self-perception, all the groups except the undergraduates attending an expensive private college reported somewhat self-achieving self-perceptions, with the artists' mean of +8.6 the strongest. It is these data, where nine of the 10 separate means in Table 1 are within 8 points from the zero midpoint, that support the conclusion that the respondents had self-perceptions with elements of both extrafamily, self-achieving, and intrafamily, other-achieving strivings, somewhat overbalanced toward the active or self-achieving satisfactions. It is well to note that every respondent studied in her self-perception had elements of both group and individual strivings, and no respondent ever scored within 20 points of the maximum, plus or minus 68. The standard deviations, of course, reflect the fact that there was considerable variability within this generalization.

The ideal woman delineated by these female respondents was slightly more self-achieving than the self-perception for seven of the nine samples where data were available, and so indicate that the composite ideal woman tends to be even more balanced than the self-perceptions between self-activating and other-activating strivings. Typically, the discrepancy between self-perception and ideal woman is about 2 points, with the largest discrepancy evidenced in the sample of Negro professionals. Here, the discrepancy was 5.8 points.

The dramatic shift in the data occurs in the women's perception of man's ideal woman. Here all the samples delineated a *strongly* family-oriented woman, considerably more so in every instance than either their own self-perceptions or their perceptions of ideal woman. In 10 of the 11 samples, only a few respondents indicated an active woman when speaking of man's ideal. Thus, on the average, while female respondents delineated a somewhat self-achieving self-perception and a slightly less self-achieving ideal woman, they attributed to men a strongly family-oriented ideal woman.

Table 1 also presents the new data on men's perception of their ideal woman. It is obvious that all six samples of men delineated an ideal woman relatively balanced between active and permissive elements.[2] Four of the samples were slightly on the active side of zero, and the other two slightly on the opposite side of zero. The discrepancy between the ideal woman men say they want and the ideal woman that women attribute to them is consistent in every comparison (ranging from 12.6 points for members of philanthropic organizations to 29.6 for public college undergraduates), and is always statistically significant at the .01 level. Moreover, the ideal woman men delineate is not significantly different from the women's own self-perceptions.

When the items themselves are studied in a second level of analysis and the female ideal and the ideal woman described by men compared, there was the same[3] response pattern for 29 of the 34 items. Minor disagreement occurred in five items; of these, four were in the personal characteristic area, and only one item might be called a male-female item: whether it was woman's responsibility to make marriage work. Two-thirds of the men said that their ideal woman would not believe it was woman's responsibility to make marriage work. The women, however, were divided on this issue. Half of the women said that their ideal woman *would* believe it was woman's responsibility, the other half noted that their ideal woman did not think that a working marriage *was* woman's responsibility.

---

[2] **It is important to note that this discussion is considering relative positions on a continuum, not absolutes.**

Table 2 presents data on the similarity of response pattern on what women *thought* men wanted in an ideal woman and man's expressed concept of his ideal woman. In contrast to the agreement on 29 of the 34 items when each sex described its own ideal woman, this comparison of item response patterns for male ideal and the ideal woman attributed to them indicates the same response pattern for only 15 items of the 34. Thus there was a different response pattern for the remaining 19 items. Table 2 organizes the items into the four major areas noted earlier. Within each area, the items presented first are those in which the response pattern for male and female was the same. Those for which the response pattern was different are listed next.

Throughout Table 2, the letter "A" is used to represent a pattern in which the majority of the respondents agreed with the item. The letter "D" is used when the majority of the respondents disagreed with the item. The letter "B" is used to indicate what the authors consider the bimodal response, in which a majority of the respondents neither agreed nor disagreed. For each pattern, the proportions are also given.

In an examination of Table 2, in Section I—Attitude Toward Work and Accomplishment, one notes that both sexes agreed that man's ideal woman would prefer the constant affection of one man to admiration throughout the nation, and would believe that one's greatest satisfactions are obtained through one's own efforts. For the five other items that had different response patterns, it is of interest that on all five there *was* a modal response from the males, but for females there was a bimodal response for three, and a weak modal response for the other two. Thus, females as a group had no clear response pattern to these items.

Section II includes the largest cluster of items, 14, on Attitude Toward Marriage. For half of the items, men and women had the same response

---

[3] Response patterns were considered the "same" when they had the same modal response; they were considered "different" when they had different modal responses.

TABLE 2
MODAL RESPONSE PATTERN FOR MAN'S IDEAL WOMAN, BY SEX, MODAL PATTERN, AND PER CENT

		Same response pattern		Different response pattern			
		Male	Female	Male		Female	
Paraphrase of item	Type[a]	%	%	Type	%	Type	%
I. Attitude Toward Work and Accomplishment as Woman, Wife, and Mother							
1. Would rather be famous than have affection*	D	81	89				
2. One attains satisfactions through one's own efforts	A	76	51				
3. Like to create something**				A	76	B	43-35
4. Capable woman has duty to be active				A	71	B	49-37
5. Woman who works cannot be as good a mother**				D	67	A	51
6. Unattractive women are most ambitious				A	55	B	47-35
7. Working mother can be as good a mother***				A	52	D	55
II. Attitude Toward Marriage both as a Goal for the Single Woman, and in Terms of the Husband-Wife Relationship							
8. Like to marry a man I look up to	A	80	91				
9. Encouragement greatest contribution of wife	A	68	78				
10. Conflict between fulfillment as woman and individual	D	68	60				
11. Better to marry person below her ideal than remain unmarried	D	68	51				
12. Ambitions should be subordinated to family	A	64	82				
13. Rather not marry than sacrifice beliefs	D	61	67				

In Section III, Attitude Toward Child-Rearing, two more concepts were added to the intrafamily cluster on which both men and women agreed: that the main goal of a woman's life is to raise well adjusted children, and that the joys of motherhood make up for the sacrifices. Of interest for future generations is the fact that men saw their ideal woman as rearing children to believe in the equality of the sexes.

Section IV, with items dealing with Characteristics of the Self, includes references toward attitudes concerning authority and leadership, and two of them point up the paradox in much of these data: both men and women agreed that a woman should not only be capable of putting herself in the background and working for a person whom she admires, but also that she should be energetic in the development and expression of her ideas. This contradiction in attitude reveals, at least, that both the men and women agreed in their ambivalent attitude.

E. Conclusions and Implications

What conclusions may be drawn from these data? It appears that certain rather definite statements may be made within the limitations of both instrument and samples. In examining Table 1, one pattern in the data is very clear: in all samples, these American women said "I am pretty much what I would like to be, but I am not what men would like me to be." This pattern, moreover, is one wherein women attributed to men a strongly family-oriented ideal woman. On the other hand, American men, when questioned, stated that their ideal woman had a balance of intrafamily and extrafamily feelings. The discrepancy between the family-oriented, permissive woman that women believed men desired and the ideal woman that men actually delineated may be accounted for in at least four ways. In responding as they do: (*a*) both men and women were telling the truth, and the data reflect serious lack of communication; (*b*) women were projecting what they really might be feeling and what they would like men to believe; (*c*) men were *talking* a current liberal stereotype and may neither believe nor behave as they talked; and (*d*) when the men espoused a liberal concept of woman's role, they were exposing an ambivalence that may or may not be the result of their own use of projection and denial of an active role.

New research now underway is intended to test which of these four interpretations is the most sensible explanation of the discrepancy. However, at this time it is possible to state that, in the eyes of these American women, American men desired a type of woman that women had no wish to be. The discrepancy between ideas is so great, the gap is so wide and deep that even more than a lack of understanding this discrepancy suggests a real lack of communication between men and women. Evidently both men and women do not understand each one's desires as to what role a woman should assume.

In Table 2, in an examination of the specific items, certain points stand out clearly. First, men, in many respects, *were* delineating a more liberal point of view than the women attributed to them; but second, there was an internal contradiction in the responses of the men, a conceptual ambivalence.

Consider first the men's "liberal" responses. For example, at least 70 per cent of the men said their ideal woman would believe that a capable woman has the duty to be active and that their ideal woman would like to create something important, and also that a wife's opinion is as important as the husband's opinion. In each of these instances, the women did not think that a man's ideal woman would think this way. There is real conflict within these perceptions.

But was the men's "liberal" perception of an ideal woman consistently liberal? Actually, the liberal expressions seem to be elicited by the global items. The men said that a woman should be "active" outside the family, have "responsibility," use her "talents," "create," "fulfill" herself. However, when men were asked "What do you think about marriage and children as the most important aspects of a woman's life?" the men's liberality was not so clear. They split, a clear note of ambiguity. One might hypothesize that on a generalization level, men did accept woman's avenues of fulfillment, but when they were asked "Does this mean that woman's own work, woman's own self-realization, might be at some time the most important aspect of her life?" these men answered "No."

The women, in a sense, may have been more realistic than the men. They did not go along even with generalities, although they were split on some. Note the bimodal responses in Section II to items on whether a capable woman has a duty to be active or to create something important or whether attractive women are more ambitious than unattractive women. Here we see ambivalence too, within the women's perception of men's ideal woman.

But this ambivalence seems different from what we have described as the internal contradiction ambivalence of men. As a group men gave clear modes, but these modes often conflicted in concept: i.e., the liberal position men took on global items *versus* the relatively traditional position they took on specific items. In contrast, the women were ambivalent as a group, reflected in the bimodal responses just cited. This group ambivalence suggests that the American women studied often did not have any clear perception of what men want in a woman. But if men *are* paradoxical and *are* feeding contradictory cues to women, obviously the women would not know how men think, which would give rise to the bimodal pattern for women. The data suggest men are saying to women, "Yes, you are very bright and with this brightness you should do something," but at the point when the women stir themselves and do "something," the men add "but you can't do what you're doing because of the kids." The woman then says to herself, "What does this man want me to do?" Obviously, she is going to be uncertain.

Perhaps there is a suggestion here that at this time, in this "gray new world" (2) both men and women are rejecting each one's desires as to what role they should assume.[4]

REFERENCES

1. STEINMANN, A., LEVI, J., & FOX, D. J. Self-concept of college women compared with their concept of ideal woman and men's ideal woman. *J. Counsel. Psychol.*, 1964, **2**(4), 370-374. Read at the American Psychological Association Convention, Philadelphia, Pennsylvania, August, 1963.

2. WICKER, T. A melancholy link to yesteryear. *New York Times*, August 31, 1966.

9 East 81st Street
New York, New York 10028

---

[4] This paper was read at the Eastern Psychological Association in New York, April 16, 1966.

Published as a separate and in *The Journal of Psychology,* 1966, **64**, 277-288.

DESENSITIZATION FACTORS IN THE HYPNOTIC TREATMENT OF PHOBIAS*

Department of Psychology, Boston College

JOSEPH R. CAUTELA

A. INTRODUCTION

The success of behavior therapy (22, 29, 39) in the treatment of neuroses has led some investigators to posit variables other than the learning variables (27) to account for successful treatment outcome. The concept of hypnosis is one of the variables. This is true especially when desensitization procedure is used. Since Wolpe has used hypnotic induction procedures to relax his patients and since the usual method based on Jacobson's (20) procedure, the question of the relation between desensitization and hypnosis seems quite appropriate. In a previous paper (8), the author explored the possibility of accounting for the effectiveness of the desensitization procedure in terms of the variables subsumed under the rubric of "hypnosis." By presenting a logical analysis of the concept of hypnosis, and by citing experimental evidence, the author attempted to demonstrate that hypnosis or suggestion is not a viable aspect of the success of the desensitization procedure.

Also in this paper (8), the author argued in the following manner: Two of the components present in the desensitization procedure also present in the hypnotic induction procedure are instructions to the *S* to close his eyes and suggestions of relaxation. Adding the other components of the hypnotic induction procedure (defining the situation as hypnosis, suggestions of sleep, suggesting to *S* that he will find it easy to respond to test suggestions, suggestions given to motivate the *S* to perform to the best of his ability) prior to desensitization does not contribute to the efficacy of the desensitization procedure (21). Also it appears that *S*'s ability to respond to test suggestion is not related to therapeutic outcome (21).

A logical analysis of the concept of hypnosis as commonly used refers indiscriminately to many independent variables subsumed under the words, "induction procedure," to a hypothetical construct (the so-called altered state of consciousness), and to many dependent variables (observable behavior that is said to be the result of the hypnotic induction procedure). An examination

---

* Received in the Editorial Office, Provincetown, Massachusetts, on September 26, 1966, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

of the relationship among these variables reveals that the indication of an altered state of consciousness is not necessary to account for the behavior subsumed under the dependent variables. Barber (4) has shown that almost all the features of so-called hypnotic behavior, such as amnesia, analgesia, hallucinations, age regression, and body rigidity, are a function of specific denotable independent variables. These variables include, for example, motivational instructions and a definition of the situation as one in which this behavior is expected. It is evident from Barber's work (2, 3, 4, 5) that it is not necessary to postulate an altered state of consciousness as an explanation of the behavior that occurs when the hypnotic induction procedure is used.

After the author investigated the possibility of trying to account for the effectiveness of the desensitization procedure in terms of hypnotic phenomena, it occurred to him to examine the other side of the coin. If there was some similarity between hypnosis and behavior therapy, perhaps one should look into the possibility that successful treatment of phobias with the use of hypnosis can be explained in terms of desensitization principles. It is the purpose of this paper to explore such a possibility.

In such an investigation it is necessary to describe what the behavior therapist says he does when he uses systematic desensitization in treating a phobia. It is also important to determine what the therapist does who uses the hypnotic induction procedure in treating phobias. The theoretical basis for these operations should also be explored. This examination will reveal that the results obtained by the hypnotherapist in treating phobias are not due to the presumed causal factor of an altered state of consciousness (hypnosis), but are due to the use of unsystematic desensitization procedures.

B. Systematic Desensitization Based on Relaxation

In the desensitization procedure as proposed by Wolpe (39), questionnaires and interviews are used to construct carefully a hierarchy concerning the patient's fears. The patient is then taught muscle relaxation based on Jacobson's (20) procedure. This takes anywhere from one to eight sessions. The patient is also told to practice at home. When the patient is able to relax sufficiently for him to report that he feels no tension whatsoever, the desensitization procedure is begun. The patient is asked to try to visualize as vividly as possible the item on the bottom of the hierarchy (this is the least fearful). If he can visualize the scene without any tension, he is asked to do nothing. If he feels tension or fear, he is instructed to raise his finger. If the patient does not raise his finger, the next item is presented on the hierarchy and so on until the top of the hierarchy is reached. Sometimes each item is presented two or three times before the therapist proceeds with the next item. There is usually a direct transfer from the patient's ability to visualize the scene without apprehension in the office to the actual scene outside the office. If an item causes apprehension, the subject is asked to relax again, and an item lower down on the hierarchy is presented.

C. THEORETICAL BASIS FOR SYSTEMATIC DESENSITIZATION BASED ON RELAXATION

It is Wolpe's contention that "If a response antagonistic to anxiety can be made to occur in the presence of anxiety-evoking stimuli so that it is accompanied by a complete or partial suppression of the anxiety responses, the bond between these stimuli and the anxiety responses will be weakened" (39, p. 71). Wolpe (39, p. 72) holds that for the most part responses that oppose anxiety largely implicate the parasympathetic nervous system. This counterconditioning model essentially states that a stimulus-response (S-R) bond can be broken by allowing the S to occur in the presence of a new R_1, while preventing the old R from occurring. A new S-R_1 will then develop.

D. HYPNOTIC METHODS USED IN THE TREATMENT OF PHOBIAS

A systematic investigation by the author of some 40 to 50 texts dealing with hypnosis and a search of the journals dealing exclusively with hypnosis reveal a surprising dearth of investigation and discussion in dealing with phobias as such. Most of the texts only devote one page or paragraph to the topic. Probably one of the reasons for the lack of space devoted to the topic is a relative lack of success in treating phobias by hypnosis. Dorcus (13) claims that phobias do not yield as readily to hypnosis as do compulsions. Meares (24) asserts that many phobic conditions are best treated by insight therapy rather than by suggestive therapy. Wolberg (37) discusses two cases of phobias in his book on hypnoanalysis. In both these cases treatment took years before the phobias were eliminated. Enough information was gained from a few texts that deal with the topic at some length and from a few journal articles devoted exclusively to the topic to classify three hypnotic procedures used in the treatment of phobias. The three methods are (*a*) direct suggestion while in a "hypnotic state," (*b*) hypnoanalysis, and (*c*) scene visualization while "hypnotized."

Before discussing the three methods it is important to note that relaxation instructions are usually an important component of the hypnotic induction procedure [see Barber (4)]. Cooke (11) places a great deal of emphasis on relaxing different muscles of the body in the hypnotic induction procedure.

In his discussions of different hypnotic induction procedures, Teitelbaum (35) frequently uses suggestions of relaxing. Christenson (9) defines hypnosis as a state of relaxed concentration. Magda Arnold (1) ascribes much of the hypnotic behavior as due to deep muscle relaxation. She refers to Jacobson's (20) work in this regard. In discussing the use of hypnosis in anesthesiology, Coppolino (12) stresses the importance of relaxation instructions in the induction procedure.

1. *The Use of Direct Suggestion*

This is probably the least used technique. Writers, such as Wolberg (38) and Schneck (31), stress the futility of treating phobias by direct suggestion. Wolberg (38, p. 233) states that a direct attack on a phobia is usually futile even when strong injunctions are given the patient while he is in a deep trance. The lack of success in treating phobias by direct suggestion is hardly surprising when poor results with this method can be predicted both from a psychoanalytic and learning theory point of view. As will be seen below, the psychoanalytic treatment of phobias places great stress on exploring unconscious material and "working through" the conflict (14, 15). Learning theorists (18, 19, 26) and behavior therapists (22, 39) stress the results of repeated exposure to the fearful stimulus which is not accompanied by the reinforcing stimulus (extinction) or which is presented in the presence of a response incompatible with fear (counter-conditioning).

Unless one ascribes some magical properties to the hypnotic induction procedure, such as an altered state of consciousness, there is no reason to expect successful treatment by direct suggestion unless it can motivate the patient to expose himself gradually to the fearful situation.

2. *The Use of Hypnoanalysis*

In order to get a comprehensive picture of desensitization factors in hypnoanalysis, it is necessary to present four major aspects of hypnoanalysis: (*a*) rationale for the use of hypnosis, (*b*) psychoanalytic explanation of phobic states, (*c*) methods of hypnoanalysis, and (*d*) hypnoanalysis and desensitization.

a. Rationale for the use of hypnoanalysis. In hypnoanalysis, the hypnotic induction procedure is used to shorten two phases of the therapeutic process (37, p. 162). The first phase is the uncovering of unconscious motivational patterns. According to Wolberg, this phase is very slow and may require years of treatment. Hypnosis may circumvent and more easily resolve ego defenses that keep unconscious material from gaining access to awareness (37,

p. 163). The second phase of therapy is considered even more prolonged than the first. The second phase is re-educative and involves an elaboration of new and adaptive interpersonal attitudes. Wolberg claims that social reintegration does not occur automatically. It is a slow reconditioning process, necessitating establishment of new habit patterns to displace old, destructive ones. Wolberg holds that hypnosis (37, p. 163) can shorten this phase of treatment, since the hypnotic relationship permits easier incorporation and absorption of wholesome interpersonal attitudes.

b. Psychoanalytic explanation of phobic states. As can be seen from the above description of hypnoanalysis as viewed by Wolberg, the chief proponent of hypnoanalysis, the procedure has the psychoanalytic model as its base.

Freud's (15) exposition of his analysis of Little Hans provides the basic psychoanalytic explanation concerning the formations and treatment of phobic states. Little Hans was afraid to go out in the street because he was fearful that a horse might bite him. Freud's analysis of Hans was made through discussions with the father. Freud never talked to the boy directly. Freud's conclusion was that there was a conflict between Hans' instinctual impulses and the demands of the ego. As a result of the oedipal conflict Hans was afraid of being castrated by the father. This fear was displaced onto an external danger that could sometimes be avoided. Treatment consisted in bringing out the repressed material and working through the conflict. Apparently even within the psychoanalytic model there are times when the above procedure is not sufficient. In Freud's (16) paper, "Turning of the ways of psychoanalytic therapy," he stresses the importance of exposing the patient to the feared situation or object. He states that one can hardly ever master a phobia if one waits until the patient lets the analysis influence him to give it up. It is often necessary for the therapist to intervene more actively when treating phobic patients. The therapist encourages the patient to encounter the phobic object or situation.

With the psychoanalytic basis of a phobia providing a background, a brief discussion of hypnoanalytic methods is in order. These methods aim at getting at unconscious conflicts and working them through.

c. Hypnoanalytic methods. In the use of hypnoanalysis the patient is usually somewhat relaxed; the degree of relaxation depending, for the most part, on the kind of hypnotic induction procedure employed. If the patient is lying down or in a comfortable chair with his eyes closed, he will probably be more relaxed than if he is put into a more active waking hypnosis in which his eyes are open, and he is instructed to act as if he were awake. Wolberg (37, p. 174) prefers the couch position, and he instructs the patient that although he is in a trance and asleep he is able to act and feel exactly as he wants to. Wolberg attempts to produce somnambulism whenever possible. In Wolberg's approach the patient's eyes are closed or open depending upon the hypnoanalytic technique employed at the time. In the dream induction procedure, for instance, the patient's eyes are usually closed. It is obvious that, for the most part, the patient is well relaxed when hypnoanalysis is employed.

During the hypnoanalysis the patient may be exposed, in imagination, to the phobic stimulus in any of the following ways:

1. The patient is asked to recall the earliest memory relating to the phobic stimulus (23, p. 219; 37, p. 236).
2. While the patient is in a medium or deep trance, he is given a suggestion that he will have a dream and a specific topic is suggested to him. Often the topic concerns the phobic object itself or some situation that the therapist feels is related to the phobic object (33, p. 90).
3. Free association is given to the patient and often the associations concern the phobic object (37, p. 176).
4. Automatic writing and drawing are used to get at repressed material (37, p. 194). Often the automatic writing concerns the feared object (if the patient is being treated for this kind of problem).
5. In the use of crystal and mirror gazing, the patient is asked to look at a mirror or crystal and describe what he sees while under a hypnotic trance. There are times when he sees situations depicting the feared object or situations. This method will be discussed in more detail later.

There are a number of other procedures, such as dramatics, regression and revivication, and the induction of experimental conflicts. If these techniques are used to treat phobias, they often involve exposure to the phobic object.

d. Hypnoanalysis and desensitization. It is quite evident that hypnoanalysis many times will involve exposure in imagination to the phobic object or situation while the patient is relaxed. This procedure differs in one very essential aspect from systematic desensitization based on relaxation. As previously indicated in systematic desensitization, a hierarchy is carefully (and often painstakingly) constructed concerning the patient's anxiety responses to various aspects of the phobic object or situation. The patient is very gradually exposed to the phobic stimulus while relaxed. It is important not to expose the patients to the phobic objects while the possibility of an anxiety response to it exists. It is especially important to avoid the presentation of a stimulus that results in a strong anxiety reaction. If this is done, it will probably result in another learning trial that will strengthen the association between the phobic stimulus and the anxiety response. In hypnoanalysis the patient is randomly exposed to the phobic object. Sometimes there will be a strong anxiety reaction and sometimes it will be minimal. According to what one would predict from behavior theory, the hypnoanalytic procedure will require many more sessions for a remission of the phobic symptoms, since sometimes the anxiety responses will be extinguished (counter-conditioned) and sometimes they will be reinforced. This indeed seems to be the case. Wolberg reports some cases that took years to achieve a remission. Lazarus (22), Paul (30), and Wolpe (39) offer evidence that phobias can usually be treated in 10 to 20 sessions. Of course, one could claim that the comparison of the behavior therapy treatment of a phobia with the hypnoanalytic treatment is unfair, since hypnoanalysis aims at more than eliminating the symptom. It aims at understanding the symptom. One could counter this argument by pointing out that there is no evidence that indicates that the supposed understanding and insight into the symptom is the more effective kind of treatment. By more effective the author would include less possibility of relapse, and less possibility of symptom substitution. The one excellent study done in this area of comparing an insight approach with a behavior therapy approach indicated a desensitization approach to be very much superior to an insight approach in respect to degree of improvement, symptom substitution, and relapse (30). Also the author has argued elsewhere (7) that insight may be a result of desensitization rather than the cause.

3. *Scene Visualizations While Hypnotized*

This is one of the methods employed by Wolberg (37, p. 214) in hypnoanalysis, but it is also used as a specific method for the treatment of a phobia without necessarily being a part of the hypnoanalytic procedure. The use of scene visualizations employing a crystal ball or mirror is discussed by Schneck (32, 33, 34). Scene visualizations have also been employed by Mordey (25) and Naruse (28) to eliminate anxiety or fear responses to specific stimuli.

In the crystal ball technique the patient is usually instructed in one of the following three ways:

1. The patient is told to visualize a crystal ball and asked to recall the first fear experience with the phobic object (31, 34).

2. The patient is asked to look at a crystal ball and see the phobic object under different circumstances (32, p. 27).

3. The patient is asked to visualize himself approaching the phobic object, thereby gradually exposing himself to the phobic object in small amounts (31). Though not asking a patient to visualize a crystal ball, Mordey (25) and Naruse (28) essentially employed this latter procedure.

Using the above procedures Schneck claims to have cured a dog phobia, a cat phobia, and a swimming phobia (31, 32). Naruse (28) used the scene visualization procedure to treat successfully stage fright in champion athletes. Mordey (25) also used the scene visualization procedure to treat an actor of stage fright. Sometimes, of course, all three procedures are used for the same patient.

There are obvious similarities in the scene visualization procedures and systematic desensitization. The patient is relaxed and exposed a number of times to the phobic object. Of course, a hierarchy is not built up according to the degree of anxiety responses to different aspects of the phobic situation. Also the first two scene-visualization procedures, where the subject is asked to recall his first experience or see something concerning the phobic object, are procedures that the behavior therapist avoids because they will probably result in some anxiety response in the presence of the imagined phobic object. This probably will not lead to reciprocal inhibition, but will be another trial of pairing the phobic object (S) with anxiety (R), which is apt to be reinforced by avoidance when the patient stops visualizing the scene. The third type of scene visualization, wherein the subject is asked to expose himself gradually to the phobic object, most nearly resembles the systematic desensitization procedure. It differs from it in that the gradation is not carefully developed with cooperation of the subject.

E. Discussion

It is clear that in some cases investigators employing hypnotic induction procedures have been able to treat certain kinds of phobic states successfully. It is also clear that systematic desensitization based on relaxation has been used quite successfully to treat phobias. Is there a common element (or elements) that can be used to explain the success of the two methods?

One of the common features of the two approaches is that they differ to a great extent from the traditional psychotherapeutic techniques as viewed by the patient. Perhaps the patient ascribes some special healing power to these procedures. It would be difficult, if not impossible, to completely rule out this feature as the main determining factor in therapeutic success, but there is some evidence to indicate that this is not the whole answer. One type of evidence is that direct suggestion alone is usually not effective in treating phobic states. Also, when hypnoanalysis is used the treatment can often take years. If the reason for the cure is the magical property assigned by the patient to the procedure, one would expect that the cure would take place early in treatment. When using the desensitization procedure, Lang

(21), in a carefully controlled study, varied the order of relaxation, presenting the order of the scenes and the presentation of the hierarchy. He found the only significant results were those obtained by the group that was relaxed first, then given the hierarchy by starting at the bottom and working up.

Another common feature could be that the behavior therapist has greater faith in his procedures than does the therapist employing more traditional methods. First of all we have no evidence indicating that this is so. Secondly, we have no sound evidence on how therapist expectancy influences therapeutic results (17).

Is the reason for the success of the two procedures due to the fact that the patient is in a special altered state of consciousness (hypnotized)? As previously mentioned, the author has argued elsewhere (8) that a logical analysis of the concept of hypnosis and experimental evidence indicate that this explanation lacks validity.

Is the crucial similarity in hypnosis and systematic desensitization based on relaxation? It appears to be that both employ a kind of relaxation and repeated exposure to the stimulus situation. In both cases reciprocal inhibition is involved [or "counter-conditioning," as Wolpe (39) explains reciprocal inhibition]. Of course the author may have committed a sin of omission by not including all of the logically possible similarities in the two approaches, but the number of logically possible explanations of behavior that occur in any dyadic relationship is very great. One can only pick those variables that seem most relevant, as indicated by observation.

It is interesting to note that a number of investigators other than behavior therapists also emphasize repeated exposure to the phobic object or the importance of presenting reciprocally inhibiting factors in the treatment of phobias. Of course most hypnotists who treat phobias employ repeated exposure. As mentioned earlier, Freud thought exposure to the phobic object essential in the treatment of a phobia. In an article concerned with the treatment of stage fright, Naruse (28, p. 68) actually uses the term "desensitize." In his discussion concerning the treatment of an individual with the fear of audiences, Naruse states: "A deep hypnotic trance was involved, and he was trained to visualize vividly a scene at a game in order to *desensitize* his fear or anxiety of large audiences and to familiarize him with the playing field."

Ulett and Peterson (36) state that hypnosis combined with techniques of reciprocal inhibition may be useful in the treatment of phobias. Bernstein (6) treated burned children by the use of hypnosis. He held their hands and reassured them, while using the induction procedure. Suggestions were given that they feel at ease about the things they were afraid of.

F. Further Implications

It is quite likely that the therapeutic results obtained in the treatment of maladaptive behavior other than phobias can be fruitfully explained within a behavior therapy model of either reciprocal inhibition or operant conditioning. This, of course, has to be demonstrated. It is to the credit of the hypnotherapist that he states what he does more carefully and more clearly than other more traditional therapists. It would be much more difficult to apply this type of analysis to the usual psychoanalytic brand of therapy because it is so hard to figure out what they actually do.

In writing this kind of paper, one runs the risk of hearing "Here we go again. Somebody is trying to explain everything else in terms of his own pet theory. He is no different from the psychoanalyst who feels he can explain behavior therapy (and every other kind of therapy) in terms of the psychoanalytic model." If this can be done fruitfully, it should be done. Some have attempted to explain the results of behavior therapy in terms of the phenomena of suggestion (27), and the author has written a reply to such an attempt (8).

One final point concerning hypnosis: Conant (10) once said that contradictory facts never eliminate a theory. Only when there is an alternate theory can the old theory be replaced. This is quite true in dealing with hypnosis. No matter how much evidence investigators, such as Barber (2, 3, 4, 5), present to indicate that hypnotic phenomena can be adequately explained without positing "an altered state of consciousness" as an explanation, hypnotherapists will point to the effectiveness of the procedure. This paper has been an attempt to offer an explanation other than that of an "altered state of consciousness" to explain some aspects of hypnotic phenomena.

References

1. Arnold, M. Brain function in hypnosis. *Internat. J. Clin. & Exper. Hypn.*, 1959, **7**, 109-119.
2. Barber, T. X. The effects of "hypnosis" on pain: A critical review of experimental and clinical findings. *Psychosomat. Med.*, 1963, **25**, 303-333.
3. ———. "Hypnosis" as a causal variable in present-day psychology: A critical analysis. *Psychol. Rep.*, 1964, **14**, 839-842.
4. ———. Experimental analysis of "hypnotic" behavior: A review of recent empirical findings. *J. Abn. Psychol.*, 1965, **70**, 132-154.
5. Barber, T. X., & Calverley, D. S. "Hypnotic behavior" as a function of task motivation. *J. of Psychol.*, 1962, **54**, 363-389.
6. Bernstein, M. R. Observations on the use of hypnosis with burned children on a pediatric ward. *Internat. J. Clin. & Exper. Hypn.*, 1965, **13**, 1-10.
7. Cautela, J. R. Desensitization and insight. *Behav. Res. Ther.*, 1965, **3**, 59-64.
8. ———. Hypnosis and behavior therapy. *Behav. Res. Ther.*, 1966, **4**, 219-224.
9. Christenson, J. A. An operational approach to hypnosis. *Internat. J. Clin. & Exper. Hypn.*, 1956, **4**, 89-91.

10. CONANT, J. B. On Understanding Science. New York: New Amer. Library, 1953.
11. COOKE, C. E., & VAN VOGT, A. E. The Hypnotism Handbook. Los Angeles, Calif.: Griggin, 1956.
12. COPPOLINO, C. A. Practice of Hypnosis in Anesthesiology. New York: Grune & Stratton, 1965.
13. DORCUS, R. M. The influence of hypnosis on learning and habit modifying. In R. M. Dorcus (Ed.), *Hypnosis and Its Therapeutic Applications* (5th ed.). New York: McGraw-Hill, 1956. Pp. 1-15.
14. FRIEDMAN, P. The phobias. In S. Arteti (Ed.), *American Handbook of Psychiatry (Vol. 1)*. New York: Basic Books, 1959. Pp. 293-306.
15. FREUD, S. Analysis of a phobia of a five-year-old boy. In E. Jones (Ed.), *Collected Papers (Vol. 3)*. New York: Basic Books, 1959. Pp. 149-289.
16. ———. Turning of the ways of psychoanalytic therapy. In E. Jones (Ed.), *Collected Papers (Vol. 2)*. New York: Basic Books, 1959. Pp. 392-402.
17. GOLDSTEIN, A. P. Therapist-Patient Expectancies in Psychotherapy. New York: Pergamon Press, 1962.
18. GUTHRIE, E. R. The Psychology of Human Learning. New York: Harper, 1935.
19. HULL, C. L. Principles of Behavior. New York: Appleton, 1943.
20. JACOBSON, E. Progressive Relaxation. Chicago, Ill.: Univ. Chicago Press, 1938.
21. LANG, P. The mechanism of fear reduction: A laboratory analysis of desensitization therapy. Paper read at the American Psychological Association, Chicago, Illinois, September 6, 1965, as part of the Symposium on Behavior Therapies: An Evaluation.
22. LAZARUS, A. A. The results of behavior therapy in 126 cases of severe neurosis. *Behav. Res. Ther.*, 1963, **1**, 69-79.
23. LECRON, L. M., & BORDEAUX, J. Hypnotism Today. New York: Grune & Stratton, 1949.
24. MEARES, A. A System of Medical Hypnosis. Philadelphia, Pa.: Saunders Co., 1961.
25. MORDEY, T. Conditioning of appropriate behavior to anxiety producing stimuli: Hypnotherapy of a fright case. *Amer. J. Clin. Hypn.*, 1965, **2**, 117-121.
26. MOWRER, O. H. Anxiety reduction and learning. *J. Exper. Psychol.*, 1940, **27**, 497-516.
27. MURRAY, E. J. Learning theory and psychotherapy: Biotropic versus sociotropic approaches. *J. Counsel. Psychol.*, 1963, **10**, 250-255.
28. NARUSE, G. Hypnotic treatment of stage fright in champion athletes. *Intern. J. Clin. & Exper. Hypn.*, 1965, **13**, 63-70.
29. PAUL, G. Effects of insight, desensitization, and attention placebo treatment of anxiety: An approach to outcome research in psychotherapy. University Microfilms, Ann Arbor, Michigan, 1964.
30. ———. Comparative therapeutic outcome in controlled research. Paper read at the American Psychiatric Association meetings in Atlantic City, New Jersey, May, 1966.
31. SCHNECK, J. M. The hypnoanalysis of phobic reactions. In L. M. LeCron (Ed.), *Experimental Hypnosis*. New York: MacMillan, 1954. Pp. 465-476.
32. ———. Studies in Scientific Hypnosis. Baltimore, Md.: Williams & Wilkins, 1954.
33. ———. Clinical and experimental aspects of hypnotic dreams. In M. V. Kline (Ed.), *Clinical Correlations of Experimental Hypnosis*. Springfield, Ill.: C C Thomas, 1963. Pp. 75-100.
34. ———. Hypnosis in psychiatry. In J. M. Schneck (Ed.), *Hypnosis in Modern Medicine*. Springfield, Ill.: C C Thomas, 1963. Pp. 169-203.

35. TEITELBAUM, M. Hypnotic Induction Techniques. Springfield, Ill.: C C Thomas, 1965.
36. ULETT, G. A., & PETERSON, D. B. Applied Hypnosis and Positive Suggestion. St. Louis, Mo.: Mosby, 1965.
37. WOLBERG, L. R. Hypnoanalysis. New York: Grune & Stratton, 1945.
38. ———. Medical Hypnosis. New York: Grune & Stratton, 1948.
39. WOLPE, J. Psychotherapy by Reciprocal Inhibition. Stanford, Calif.: Stanford Univ. Press, 1958.

Department of Psychology
Boston College
Chestnut Hill, Massachusetts 02167